

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 18, 1999, 14:29:32 ; Search time 1261.22 Seconds
(without alignments)
12184.436 Million cell updates/sec

Title: US-09-001-039A-46
Perfect score: 4832
Sequence: 1 CTCGAGCTAAGATATTTTA.....ATTATGATGTGCGGCCGC 4832

Scoring table: IDENTITY_NUC

Searched: 679419 seqs, 1590154680 residues

Database :

- GenEmbl: *
- 1: gb_ba1:*
 - 2: gb_ba2:*
 - 3: gb_om:*
 - 4: gb_ov:*
 - 5: gb_pat:*
 - 6: gb_ph:*
 - 7: gb_pl1:*
 - 8: gb_pl2:*
 - 9: gb_pr1:*
 - 10: gb_pr2:*
 - 11: gb_pr3:*
 - 12: gb_ro:*
 - 13: gb_st:*
 - 14: gb_sts:*
 - 15: gb_sy:*
 - 16: gb_un:*
 - 17: gb_vi:*
 - 18: em_fun:*
 - 19: em_htg:*
 - 20: em_hum1:*
 - 21: em_hum2:*
 - 22: em_in:*
 - 23: em_om:*
 - 24: em_or:*
 - 25: em_ov:*
 - 26: em_pat:*
 - 27: em_ph:*
 - 28: em_pl:*
 - 29: em_ro:*
 - 30: em_sts:*
 - 31: em_sy:*
 - 32: em_un:*
 - 33: em_vi:*
 - 34: gb_htg1:*
 - 35: gb_htg2:*
 - 36: gb_in1:*
 - 37: gb_in2:*
 - 38: em_ba1:*
 - 39: em_ba2:*
 - 40: em_hum3:*
 - 41: em_hum4:*
 - 42: gb_pr4:*

BEST AVAILABLE COPY

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4570.8	94.6	9354	5	AR003585	AR003585 Sequence

2	4192.4	86.8	4278	5	I08644	I08644 Sequence 4
3	4190.8	86.7	4281	5	I08643	I08643 Sequence 3
4	4045.8	83.7	4551	5	I08641	I08641 Sequence 1
5	4045.8	83.7	4548	5	I08642	I08642 Sequence 2
6	2468.8	51.1	9009	5	AR003710	AR003710 Sequence
7	2468.8	51.1	9009	5	I31901	I31901 Sequence 3
8	2468.8	51.1	9009	5	I63424	I63424 Sequence 1
9	2468.8	51.1	8967	5	I71409	I71409 Sequence 1
10	2468.8	51.1	9009	9	HUMFVIIIT	K01740 Human coagu
11	2468.8	51.1	9029	9	HUMFVIIIT	M4113 Human coagu
12	2468.8	51.1	8967	10	HSFVIIIT	X01179 Human mRNA
13	2467.2	51.1	8241	5	A05328	A05328 Synthetic F
14	2467.2	51.1	8241	5	A07042	A07042 Artificial
15	2363	48.9	8831	5	E00527	E00527 Complete CD
16	2353.4	48.7	7440	5	I08345	I08345 Sequence 2
17	2338.4	48.4	7440	5	I08457	I08457 Sequence 1
18	2304.2	47.7	7272	5	I05404	I05404 Sequence 3
19	2293.4	47.5	7056	5	I27063	I27063 Sequence 1
20	2238	46.3	6999	5	I77105	I77105 Sequence 1
21	1813.6	37.5	7032	3	AF016234	AF016234 Canis fam
22	1806.4	37.4	7493	5	AR003712	AR003712 Sequence
23	1806.4	37.4	7493	5	I63427	I63427 Sequence 7
24	1806.4	37.4	7493	12	MUSCIVIT	L05573 Mus domesti
25	1805.6	37.4	7145	3	AF049489	AF049489 Canis fam
26	1754.8	36.3	6539	3	SSU49517	U49517 Sus scrofa
27	1723.2	35.7	1993	5	I02047	I02047 Sequence 1
28	1603.2	33.2	1728	5	I02054	I02054 Sequence 8
29	1240.4	25.7	1623	5	I27064	I27064 Sequence 3
30	1007.4	20.8	2493	9	HUMF8C	M90707 Homo sapien
31	996.8	20.6	3852	5	I04400	I04400 Sequence 35
32	996.2	20.6	3852	5	E00422	E00422 Human genom
33	879.6	18.2	1130	5	AR003711	AR003711 Sequence
34	879.6	18.2	1130	5	I31900	I31900 Sequence 1
35	879.6	18.2	1130	5	I63426	I63426 Sequence 5
36	675.4	14.0	955	5	E00526	E00526 Part of PES
37	537	11.1	2709	9	HUMFAC21	M88648 H.sapiens f
38	537	11.1	5348	10	HSAL31818	AU131818 Homo sapi
39	423.8	8.8	1728	5	I02049	I02049 Sequence 3
40	313.6	6.5	6910	3	BOVFACV	M81440 Bos taurus
41	313.6	6.5	6895	3	BOVFACV2A	M81441 Bos taurus
42	303.4	6.3	6893	10	HUMFVA	M14335 Human coagu
43	301.8	6.2	6909	5	A46255	A46255 Sequence 1
44	301.8	6.2	6909	5	A63218	A63218 Sequence 1
45	301.8	6.2	6909	9	HUMFAV	M16967 Human coagu

ALIGNMENTS

RESULT 1	AR003585	9354 bp	DNA	PAT	04-DEC-1998
LOCUS	AR003585	Sequence 2 from patent US 5744326.			
DEFINITION	AR003585				
ACCESSION	AR003585				
NID	93964844				
VERSION	AR003585.1	GI:3964844			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 9354)				
AUTHORS	Ill,C.R. and Bidlingmaier,S.				
TITLE	Use of viral CIS-acting post-transcriptional regulatory sequences to increase expression of intronless genes containing near-consensus splice sites				
JOURNAL	Patent: US 5744326-A 2 28-APR-1998;				
FEATURES	Location/Qualifiers				
source	1..9354				
BASE COUNT	2506 a	/organism="unknown"	2448 t		
ORIGIN	2239 c	2161 g			

SCHWARTZMAN
09001039
SEQ ID: 46 & 47

Query Match 94.68; Score 4570.8; DB 5; Length 9354;
Best Local Similarity 98.88; Pred. No. 0;
Matches 4643; Conservative 0; Mismatches 12; Indels 45; Gaps 2;

OY 14 TATTTAGAGAAGATTAACTTTTGTCTCTCCAGTTGAACATTTGTAGCAATAAG---T 70
|||||
Db 2904 TATTTAGAGAAGATTAACTTTTGTCTCTCCAGTTGAACATTTGTAGCAATAAGCCAC 2963

OY 71 CATGCAATAGAGCTCTCCACCTGCTTCTTCTGTGCGCTTTGCGATTGCTTTAGTGC 130
|||||
Db 2964 CATGTTATAGAGCTCTCCACCTGCTTCTTCTGTGCGCTTTGCGATTGCTTTAGTGC 3023

OY 131 CACCAGAAATACTACTCCGTGGGTGCAGTGGAACTGTCAATGGCACTATATGCAAGTGAAT 190
|||||
Db 3024 CACCAGAAATACTACTCCGTGGGTGCAGTGGAACTGTCAATGGCACTATATGCAAGTGAAT 3083

OY 191 CGGTGAGCTGCTGTGAGCCAAAGATTTCCTCTAGAGTGGCCAAATCTTTCCATTCAA 250
|||||
Db 3084 CGGTGAGCTGCTGTGAGCCAAAGATTTCCTCTAGAGTGGCCAAATCTTTCCATTCAA 3143

OY 251 CACCTCAGTCTGTACAAAAAGACTCTGTTGTAGAAATTCACGGATCACCTTTTCAACAT 310
|||||
Db 3144 CACCTCAGTCTGTACAAAAAGACTCTGTTGTAGAAATTCACGGATCACCTTTTCAACAT 3203

OY 311 CGCTAAGCCAAAGCCACCCCTGGATGGGTCTGTAGTCTTACCATCCAGCTGAGGTTTA 370
|||||
Db 3204 CGCTAAGCCAAAGCCACCCCTGGATGGGTCTGTAGTCTTACCATCCAGCTGAGGTTTA 3263

OY 371 TGATACAGTGGTCATTACACTTAAGAAACATGGCTTCCCATCTGTCACTTCATGCTGT 430
|||||
Db 3264 TGATACAGTGGTCATTACACTTAAGAAACATGGCTTCCCATCTGTCACTTCATGCTGT 3323

OY 431 TGGTGTATCTTACTGAAAGCTTCTGAGGAGCTGAATATGATGATCAGACCAGTCAAG 490
|||||
Db 3324 TGGTGTATCTTACTGAAAGCTTCTGAGGAGCTGAATATGATGATCAGACCAGTCAAG 3383

OY 491 GGAGAAAGAAATGATTAAGTCTTCCCTGGTGAAGCCATACATATGCTGGCAGTCTCT 550
|||||
Db 3384 GGAGAAAGAAATGATTAAGTCTTCCCTGGTGAAGCCATACATATGCTGGCAGTCTCT 3443

OY 551 GAAAGAGATGGTCCAAATGGCTCTGACCCACTGTGCTTACTACTCATATCTTTCTCA 610
|||||
Db 3444 GAAAGAGATGGTCCAAATGGCTCTGACCCACTGTGCTTACTACTCATATCTTTCTCA 3503

OY 611 TGTGAGCTGTAAAAAGACTTGAATTCAGGCTCATTTGGAGCCCTACTAGTATGAGAGA 670
|||||
Db 3504 TGTGAGCTGTAAAAAGACTTGAATTCAGGCTCATTTGGAGCCCTACTAGTATGAGAGA 3563

OY 671 AGGAGTCTGGCCAAAGAAAGACACAGACCTTGACAAATTTATACTACTTTTGTCTGT 730
|||||
Db 3564 AGGAGTCTGGCCAAAGAAAGACACAGACCTTGACAAATTTATACTACTTTTGTCTGT 3623

OY 731 ATTGTGAAGGAAAAAGTTGGCACTCAGAAACAAAGAACTCCTTGATGCAAGATAGGA 790
|||||
Db 3624 ATTGTGAAGGAAAAAGTTGGCACTCAGAAACAAAGAACTCCTTGATGCAAGATAGGA 3683

OY 791 TGCTGATCTGCTCGGCTGTGCTTAAATGACACAGATCAATGTTATGTAACAGGTC 850
|||||
Db 3684 TGCTGATCTGCTCGGCTGTGCTTAAATGACACAGATCAATGTTATGTAACAGGTC 3743

OY 851 TCTGCAGGTCTGATTTGATGCCACAGAAATCAGTCTATTGGCATGATTTGAATGGG 910
|||||
Db 3744 TCTGCAGGTCTGATTTGATGCCACAGAAATCAGTCTATTGGCATGATTTGAATGGG 3803

OY 911 CACCACTCTGAAGTGCATCAATATCTCTGAAGTGCACACATTTCTTTGTGAGAACCA 970
|||||
Db 3804 CACCACTCTGAAGTGCATCAATATCTCTGAAGTGCACACATTTCTTTGTGAGAACCA 3863

OY 971 TCGCCAGGCGTCTTGGAATCTGCGCAATAACTTTCCTTACTGCTCAACACTCTTGAT 1030
|||||
Db 3864 TCGCCAGGCGTCTTGGAATCTGCGCAATAACTTTCCTTACTGCTCAACACTCTTGAT 3923

OY 1031 GGACCTTGACAGTTTCTACTGTTTGTGCATATCTTCCCAACCAATGATGCGATGGA 1090

|||||
Db 3924 GGACCTTGACAGTTTCTACTGTTTGTGCATATCTTCCACCAACATGATGGCATGGA 3983

OY 1091 AGCTTATGTCAAAGTAGACAGCTGTCCAGAGAGAAACCCCACTACGAATGAAAAATATGA 1150
|||||
Db 3984 AGCTTATGTCAAAGTAGACAGCTGTCCAGAGAGAAACCCCACTACGAATGAAAAATATGA 4043

OY 1151 AGAAGCGAAAGACTATGATGATGATCTTACTGATTTCTGAATGGAATGTGTGAGTTTGA 1210
|||||
Db 4044 AGAAGCGAAAGACTATGATGATGATCTTACTGATTTCTGAATGGAATGTGTGAGTTTGA 4103

OY 1211 TGATGCAACTCTCTCTCTTCTTATCCAAATTCGCTCAGTTGCCAAGAAACATCTTAAAC 1270
|||||
Db 4104 TGATGCAACTCTCTCTCTTATATCCAAATTCGCTCAGTTGCCAAGAAACATCTTAAAC 4163

OY 1271 TTGGGTACATTACATTGCTGCTGTAAGAGAGGAGACTGGGACTATAGCTCCCTTAGTCTCGC 1330
|||||
Db 4164 TTGGGTACATTACATTGCTGCTGTAAGAGAGGAGACTGGGACTATAGCTCCCTTAGTCTCGC 4223

OY 1331 CCCCAGTACAGAAAGTTATAAAAAGTCAATATTTGAACAATGGCCCTCAGCGGATGGTAG 1390
|||||
Db 4224 CCCCAGTACAGAAAGTTATAAAAAGTCAATATTTGAACAATGGCCCTCAGCGGATGGTAG 4283

OY 1391 GAAGTACAAAAAAGTCCGATTTATGSCATACACAGATGAACCTTTAAGACTCGTAAGC 1450
|||||
Db 4284 GAAGTACAAAAAAGTCCGATTTATGSCATACACAGATGAACCTTTAAGACTCGTAAGC 4343

OY 1451 TATTCAGCATGAATCAGGAATCTTGGGACCTTTACTTTATGGGAAAGTTGAGACACACT 1510
|||||
Db 4344 TATTCAGCATGAATCAGGAATCTTGGGACCTTTACTTTATGGGAAAGTTGAGACACACT 4403

OY 1511 GTTGATTATATTAAAGATCAAGCAAGCAGACCATATTAACATCTAACCTCAGCGAATCAC 1570
|||||
Db 4404 GTTGATTATATTAAAGATCAAGCAAGCAGACCATATTAACATCTAACCTCAGCGAATCAC 4463

OY 1571 TGATGTCGCTCTTGTATTTCAAGAGAGATTACCAAAAAGGTGTAACATTTGAAGGATTT 1630
|||||
Db 4464 TGATGTCGCTCTTGTATTTCAAGAGAGATTACCAAAAAGGTGTAACATTTGAAGGATTT 4523

OY 1631 TCCAATTCTGCCAGAGAAATATTTCAAAATATAAATGACAGTGACTGTAGAAGATGGGCC 1690
|||||
Db 4524 TCCAATTCTGCCAGAGAAATATTTCAAAATATAAATGACAGTGACTGTAGAAGATGGGCC 4583

OY 1691 AACTAAATCAGATCCTCGGTGCCGTGACCCGCTATTACTCTAGTTTCGTTAATATGAGAG 1750
|||||
Db 4584 AACTAAATCAGATCCTCGGTGCCGTGACCCGCTATTACTCTAGTTTCGTTAATATGAGAG 4643

OY 1751 AGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGCTACAAAGAAATCTGTAGATCA 1810
|||||
Db 4644 AGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGCTACAAAGAAATCTGTAGATCA 4703

OY 1811 AAGAGAAACAGATTAATGTCAAGCAAGAGAAATGTCAATCTGTTTTCTGTATTTGATGA 1870
|||||
Db 4704 AAGAGAAACAGATTAATGTCAAGCAAGAGAAATGTCAATCTGTTTTCTGTATTTGATGA 4763

OY 1871 GAACCGAAGCTGTGTAACCTCACAGAGAATATATACAAAGCTTTCTCCCAATCCAGCTGGAGT 1930
|||||
Db 4764 GAACCGAAGCTGTGTAACCTCACAGAGAATATATACAAAGCTTTCTCCCAATCCAGCTGGAGT 4823

OY 1931 GCAGCTTGAGGATCCAGAGTTCACAGCCCTCAACATCATGACACAGCATCAATGGCTATGT 1990
|||||
Db 4824 GCAGCTTGAGGATCCAGAGTTCACAGCCCTCAACATCATGACACAGCATCAATGGCTATGT 4883

OY 1991 TTTTGTATAGTTTGCAAGTTGTCAAGTTGTTGGCATGAGGTGGCATACTGTAACATTTCTAAG 2050
|||||
Db 4884 TTTTGTATAGTTTGCAAGTTGTCAAGTTGTTGGCATGAGGTGGCATACTGTAACATTTCTAAG 4943

OY 2051 CATTTGAGCAGACACTGACTTCCCTTTCTGTCTCTCTCTGATATACCTTCAACACAA 2110
|||||
Db 4944 CATTTGAGCAGACACTGACTTCCCTTTCTGTCTCTCTCTGATATACCTTCAACACAA 5003

OY 2111 AATGCTATGAGAGACACACTCAACCTTATCCCATTTCTCAGAGAAACGTCTTCATGTC 2170
|||||

Db 5004 AATGCTCTATGAAGACACACTCACCCCTATTCCCATCTCTCAGAGAAACTGTCTTCATGTC 5063
QY 2171 GATGAAAAAACCAGGCTCTATGGATTCTGGGGTGCACAACCTCAGACTTTCGGAACAGAG 2230
Db 5064 GATGAAAAAACCAGGCTCTATGGATTCTGGGGTGCACAACCTCAGACTTTCGGAACAGAG 5123
QY 2231 CATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAGAACAACCTGGTGAATTATTACGAGA 2290
Db 5124 CATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAGAACAACCTGGTGAATTATTACGAGA 5183
QY 2291 CAGTTATGAAGATATTTCAGCATACTGCTGAGTAAAAACAATGCCATTGGAACCAAGAAG 2350
Db 5184 CAGTTATGAAGATATTTCAGCATACTGCTGAGTAAAAACAATGCCATTGGAACCAAGAAG 5243
QY 2351 CTCTCCAGAA-----CCGACC 2368
Db 5244 CTCTCCAGAAATTCAGACACCCTAGCACTAGGCAAAAGCAATTTAATGCCACCCACC 5303
QY 2369 AGTCTTGAACGCCATCAACGGGAAATTAATCTGTAATCTCTCAGTCAGATCAAGAGA 2428
Db 5304 AGTCTTGAACGCCATCAACGGGAAATTAATCTGTAATCTCTCAGTCAGATCAAGAGA 5363
QY 2429 AATGCACTATGATGATACCATATCAGTTGAATGAAGAAGAAGATTTGACATTTATGA 2488
Db 5364 AATGCACTATGATGATACCATATCAGTTGAATGAAGAAGAAGATTTGACATTTATGA 5423
QY 2489 TGAGAGTGAATAATCAGAGCCCCCGAGCTTTCAAAAGAAAAACAGACACTATTTTATTC 2548
Db 5424 TGAGAGTGAATAATCAGAGCCCCCGAGCTTTCAAAAGAAAAACAGACACTATTTTATTC 5483
QY 2549 TGCAGTGGAGAGGCTCTGGGATTATGGGATGAGTAGTCCCAACATGTTCTAAGAAACAG 2608
Db 5484 TGCAGTGGAGAGGCTCTGGGATTATGGGATGAGTAGTCCCAACATGTTCTAAGAAACAG 5543
QY 2609 GGCTCAGAGTGGCAGTGTCCCTCAGTTCAGAAGAGTGTTCAGAGAAATTACTGATGG 2668
Db 5544 GGCTCAGAGTGGCAGTGTCCCTCAGTTCAGAAGAGTGTTCAGAGAAATTACTGATGG 5603
QY 2669 CTCTTACTCAGCCCTTATACCGTGGAGAACTAAATGAACATTTGGGACTCCTGGGCC 2728
Db 5604 CTCTTACTCAGCCCTTATACCGTGGAGAACTAAATGAACATTTGGGACTCCTGGGCC 5663
QY 2729 ATATATAAAGAGCAGAGATTGAAGATATATCATGTGTAATCTTCAGAAATCAGGCCCTCG 2788
Db 5664 ATATATAAAGAGCAGAGATTGAAGATATATCATGTGTAATCTTCAGAAATCAGGCCCTCG 5723
QY 2789 TCCCTATTCCTCTATTTCTAGCCTTATTTCTTATGAGGAAGATCAGAGGCAAGAGACAGA 2848
Db 5724 TCCCTATTCCTCTATTTCTAGCCTTATTTCTTATGAGGAAGATCAGAGGCAAGAGACAGA 5783
QY 2849 ACCTAGAAAAAACTTGTCAAGCCTATGAACCAAAACTTACTTTGGAAAAGTGCACAA 2908
Db 5784 ACCTAGAAAAAACTTGTCAAGCCTATGAACCAAAACTTACTTTGGAAAAGTGCACAA 5843
QY 2909 TCATATGGCACCACCTAAAGATGAGTTTGACTGCAAAAGCCTGGGCTTATTTCTGTGATGT 2968
Db 5844 TCATATGGCACCACCTAAAGATGAGTTTGACTGCAAAAGCCTGGGCTTATTTCTGTGATGT 5903
QY 2969 TGACCTGGAAAAAGATGTGCACTCAGGCTGATTTGGACCCCTTCTGTCTGCCACACTAA 3028
Db 5904 TGACCTGGAAAAAGATGTGCACTCAGGCTGATTTGGACCCCTTCTGTCTGCCACACTAA 5963
QY 3029 CACACTGAACCCCTGCTCATGGGAGACAAGTGCAGATGACAGGAATTTGCTCTGTTTTCAC 3088
Db 5964 CACACTGAACCCCTGCTCATGGGAGACAAGTGCAGATGACAGGAATTTGCTCTGTTTTCAC 6023
QY 3089 CATCTTTGATGAGACCAAAAGCTGTGTAATCTCACTGAAATATGGAAGAAAACCTGCAGGCC 3148
Db 6024 CATCTTTGATGAGACCAAAAGCTGTGTAATCTCACTGAAATATGGAAGAAAACCTGCAGGCC 6083
QY 3149 TCCCTGCAATATCCAGATGGAAGATCCCACTTTAAAGAGAATATCGTTCATGCAAT 3208
Db 6084 TCCCTGCAATATCCAGATGGAAGATCCCACTTTAAAGAGAATATCGTTCATGCAAT 6143

QY 3209 CAATGGCTACATAATGATACACTACCTGGCTTAGTAATGGCTCAGGATCAAGGATTCG 3268
Db 6144 CAATGGCTACATAATGATACACTACCTGGCTTAGTAATGGCTCAGGATCAAGGATTCG 6203
QY 3269 ATGGTATCTGCTCAGCATGGGAGCAATGAAAACATCCATCTCTATTCATTTCAAGTGACA 3328
Db 6204 ATGGTATCTGCTCAGCATGGGAGCAATGAAAACATCCATCTCTATTCATTTCAAGTGACA 6263
QY 3329 TGTGTTCACTGTACAAAAAAGAGAGATATAAATGGCACTGTACAACTCTCTATCCAGG 3388
Db 6264 TGTGTTCACTGTACAAAAAAGAGAGATATAAATGGCACTGTACAACTCTCTATCCAGG 6323
QY 3389 TGTGTTTGAAGACAGTGGAAATGTATACCATCCAAAGCTGGAATTTGGGGGTGGAATGCC 3448
Db 6324 TGTGTTTGAAGACAGTGGAAATGTATACCATCCAAAGCTGGAATTTGGGGGTGGAATGCC 6383
QY 3449 TATGGCCGAGCATCTACATGCTGGGATGAGCACACTTTTCTGTGTACAGCAATTAAGTG 3508
Db 6384 TATGGCCGAGCATCTACATGCTGGGATGAGCACACTTTTCTGTGTACAGCAATTAAGTG 6443
QY 3509 TCAGACTCCCTGGGAATGGCTTCTGGACACATTAAGAGATTTTCAGATTACAGCTTCAGG 3568
Db 6444 TCAGACTCCCTGGGAATGGCTTCTGGACACATTAAGAGATTTTCAGATTACAGCTTCAGG 6503
QY 3569 ACAATATGACAGTGGGCCCCCAAGCTGGCCAGACTTCATTATTCGGATCAATCAATGC 3628
Db 6504 ACAATATGACAGTGGGCCCCCAAGCTGGCCAGACTTCATTATTCGGATCAATCAATGC 6563
QY 3629 CTGAGCACCAAGAGCCCTTTTCTGGATCAAGGTGATCTGTGGCACCNAATGATTA 3688
Db 6564 CTGAGCACCAAGAGCCCTTTTCTGGATCAAGGTGATCTGTGGCACCNAATGATTA 6623
QY 3689 TCACGGCATCAAGACCAGGGTGGCCTCAGAAAGTCTCAGCCCTCATCTCTCAGTT 3748
Db 6624 TCACGGCATCAAGACCAGGGTGGCCTCAGAAAGTCTCAGCCCTCATCTCTCAGTT 6683
QY 3749 TATCATCATGTATAGTCTGTATGGAGAAAGTGGCAGACTTATCGAAGAAATTCACCTGG 3808
Db 6684 TATCATCATGTATAGTCTGTATGGAGAAAGTGGCAGACTTATCGAAGAAATTCACCTGG 6743
QY 3809 AACCTTAATGCTCTTCTTTGGCAATGTGATTCATCTGGGATAAACACAATATTTTAA 3868
Db 6744 AACCTTAATGCTCTTCTTTGGCAATGTGATTCATCTGGGATAAACACAATATTTTAA 6803
QY 3869 CCCCTCAATTATTCGATACATCCGTTTGACCCCACTCATTAATAGCATTCGCAGAC 3928
Db 6804 CCCCTCAATTATTCGATACATCCGTTTGACCCCACTCATTAATAGCATTCGCAGAC 6863
QY 3929 TCTTCGATGGAGTTGATGGGCTGTGATTTAAATAGTTGCAGCATGCCATTGGAAATGA 3988
Db 6864 TCTTCGATGGAGTTGATGGGCTGTGATTTAAATAGTTGCAGCATGCCATTGGAAATGA 6923
QY 3989 GAGTAAGCAATATACAGATGCACAGATTACTGCTCATCTCTACTTTACCAATATGTTGC 4048
Db 6924 GAGTAAGCAATATACAGATGCACAGATTACTGCTCATCTCTACTTTACCAATATGTTGC 6983
QY 4049 CACCTGGTCTCTTCAAAAAGCTGCACCTTCACCTCCAAGGAGAGAGATATGCCGAGACC 4108
Db 6984 CACCTGGTCTCTTCAAAAAGCTGCACCTTCACCTCCAAGGAGAGATATGCCGAGACC 7043
QY 4109 TCAGGTGATAATCCAAAAGAGTGGCTGCAAGTGCAGTTCCAGAAGCAATGAAAGTCAAC 4168
Db 7044 TCAGGTGATAATCCAAAAGAGTGGCTGCAAGTGCAGTTCCAGAAGCAATGAAAGTCAAC 7103
QY 4169 AGGAGTAATCTACTCAGGAGTAATAATCTCTTACAGCATGTATGTGAAGAGTTCCT 4228
Db 7104 AGGAGTAATCTACTCAGGAGTAATAATCTCTTACAGCATGTATGTGAAGAGTTCCT 7163
QY 4229 CATCTCAGCAGTCAAGATGGCCATCAGTGGACTCTCTTTTTCAGAAATGGCAAGTAA 4288
Db 7164 CATCTCAGCAGTCAAGATGGCCATCAGTGGACTCTCTTTTTCAGAAATGGCAAGTAA 7223

QY 4289 GGTTTTTCAGGGAATCAAGACTCCTTCACACCTGTGTGTAAGTCTCTAGACCCACCGTT 4348
|||||
Db 7224 GGTTTTTCAGGGAATCAAGACTCCTTCACACCTGTGTGTAAGTCTCTAGACCCACCGTT 7283
QY 4349 ACTGACTCGCTACCTTCGAATTCACCCCGACAGATTGGGTGCACACGATTGCCCTGAGGAT 4408
|||||
Db 7284 ACTGACTCGCTACCTTCGAATTCACCCCGACAGATTGGGTGCACACGATTGCCCTGAGGAT 7343
QY 4409 GGAGGTTCTGGGCTGCGAGGCACAGGACCTCTACTGAGGGTGGCCACTGCAGCACCTGCC 4468
|||||
Db 7344 GGAGGTTCTGGGCTGCGAGGCACAGGACCTCTACTGAGGGTGGCCACTGCAGCACCTGCC 7403
QY 4469 ACTGCCGTACCTCTCCCTCTCAGCTCCAGGGCAGTGTCCCTCCCTGGCTTCCCTTA 4528
|||||
Db 7404 ACTGCCGTACCTCTCCCTCTCAGCTCCAGGGCAGTGTCCCTCCCTGGCTTCCCTTA 7463
QY 4529 CCTTGTGCTAAATCCTAGCAGACACTGCTTGAAGCCTCTGAATTAACCTATCATCAGT 4588
|||||
Db 7464 CCTTGTGCTAAATCCTAGCAGACACTGCTTGAAGCCTCTGAATTAACCTATCATCAGT 7523
QY 4589 CCTGCATTTCTTTGGTGGGGGGCCAGGAGGGTGCATCCAAATTAACCTTACTCTTACCTA 4648
|||||
Db 7524 CCTGCATTTCTTTGGTGGGGGGCCAGGAGGGTGCATCCAAATTAACCTTACTCTTACCTA 7583
QY 4649 TTTTCTGCAGCTGCTCCAG 4668
|||||
Db 7584 TTTTCTGCAGGGGATCTCAG 7603

RESULT 2
LOCUS 108644 4278 bp PAT 14-NOV-1994
DEFINITION Sequence 4 from Patent WO 8800831.
ACCESSION 108644
NID 9588650
VERSION 108644.1 GI:588650
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 4278)
AUTHORS Pasek, M.P.
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR
VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE
POLYPEPTIDES IN HIGH YIELDS
Patent: WO 8800831-A 4 11-FEB-1988;
Location/Qualifiers
1. 4278
JOURNAL /organism="unknown"
TITLES source
BASE COUNT 1244 a 942 c 946 g 1146 t
ORIGIN

Query Match 86.8%; Score 4192.4; DB 5; Length 4278;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 4271; Conservative 0; Mismatches 1; Indels 48; Gaps 3;

QY 129 GCCACCAAGAGATACCTGGGTGAGTGAAGTGTGATGGAGCTATATGCCAAGTGAT 188
|||||
Db 1 GCCACCAAGAGATACCTGGGTGAGTGAAGTGTGATGGAGCTATATGCCAAGTGAT 60
QY 189 CTGGGTGAGCTGCTGTGAGCAGCAAGATTCTCTCTAGAGTGCACCAAAATCTTTCCATT 248
|||||
Db 61 CTGGGTGAGCTGCTGTGAGCAGCAAGATTCTCTCTAGAGTGCACCAAAATCTTTCCATT 120
QY 249 AACACCTCAGTGTGTACAAAAGAGCTGTGTTGTAGAAATTCACGGATCACCTTTTCAAC 308
|||||
Db 121 AACACCTCAGTGTGTACAAAAGAGCTGTGTTGTAGAAATTCACGGATCACCTTTTCAAC 180
QY 309 ATCGCTAAGCCAAAGGCCACCTTGATGGTCTGCTAGTCTTACCATCCAGGCTGAGGTT 368
|||||
Db 181 ATCGCTAAGCCAAAGGCCACCTTGATGGGTTCTGTAGTCTTACCATCCAGGCTGAGGTT 240

QY 369 TATGATACAGTGGTCAATTACCTTAAGAACATGGCTCCCATCCGTGAGTCTTCAATGCT 428
|||||
Db 241 TATGATACAGTGGTCAATTACCTTAAGAACATGGCTCCCATCCGTGAGTCTTCAATGCT 300
QY 429 GTTGGTATCTCTACTGGAAGCTTCTGAGGGAGCTGAATATGATGATCAGACCAGTCAA 488
|||||
Db 301 GTTGGTATCTCTACTGGAAGCTTCTGAGGGAGCTGAATATGATGATCAGACCAGTCAA 360
QY 489 AGGAGAAGAAGATGATTAAGTCTTCCCTGTTGGAAGCCATACATATGTCTGGCAGTTC 548
|||||
Db 361 AGGAGAAGAAGATGATTAAGTCTTCCCTGTTGGAAGCCATACATATGTCTGGCAGTTC 420
QY 549 CTGAAGAAGATGTTCCATGGCTCTGACCCACATGTCCTTACCTACTCATATCTTTCT 608
|||||
Db 421 CTGAAGAAGATGTTCCATGGCTCTGACCCACATGTCCTTACCTACTCATATCTTTCT 480
QY 609 CATGTGACCTGTATAAAGACTTGAATTCAGGCTCATTTGGAGCCCTACTAGATGTAGA 668
|||||
Db 481 CATGTGACCTGTATAAAGACTTGAATTCAGGCTCATTTGGAGCCCTACTAGATGTAGA 540
QY 669 GAAGGAGTCTGGCCAAAGAAAGACACAGACCTTGACAAATTTACTACTTTTTCCT 728
|||||
Db 541 GAAGGAGTCTGGCCAAAGAAAGACACACACACCTTGACAAATTTACTACTTTTTCCT 600
QY 729 GTATTTGATGAAGGAAAGTGTGGCACTCAGAAACAAAGAACTCCTTGATGCAGAGTAGG 788
|||||
Db 601 GTATTTGATGAAGGAAAGTGTGGCACTCAGAAACAAAGAACTCCTTGATGCAGAGTAGG 660
QY 789 GATGCTGATCTGCTCGGGGCTGGCCCTAAATGCACACAGTCAATGCTTATGTAAACAGG 848
|||||
Db 661 GATGCTGATCTGCTCGGGGCTGGCCCTAAATGCACACAGTCAATGCTTATGTAAACAGG 720
QY 849 TCTCTG---CCAGTCTGATGTGATGCCACAGAAATCAGTCTATTGGCATGTGATTGGA 905
|||||
Db 721 TCTCTGCTACCAAGTCTGATGTGATGCCACAGAAATCAGTCTATTGGCATGTGATTGGA 780
QY 906 ATGGGACCACTCTCTGAAGTGCATCAATATTCTCGAAGGTGCACACATTTCTTGTGAGG 965
|||||
Db 781 ATGGGACCACTCTCTGAAGTGCATCAATATTCTCGAAGGTGCACACATTTCTTGTGAGG 840
QY 966 AACCATCGCCAGGCGCTCTTGAATCTCGCCAAATTAATTTCTTACTGCTCAACACTC 1025
|||||
Db 841 AACCATCGCCAGGCGCTCTTGAATCTCGCCAAATTAATTTCTTACTGCTCAACACTC 900
QY 1026 TTGATGACCTTGCAGAGTTTCTACTGTTTGTGATATCTCTTCCACCAACATGATGGC 1085
|||||
Db 901 TTGATGACCTTGCAGAGTTTCTACTGTTTGTGATATCTCTTCCACCAACATGATGGC 960
QY 1086 ATGGAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGGAACCCCACTACGAATGAAAAAT 1145
|||||
Db 961 ATGGAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGGAACCCCACTACGAATGAAAAAT 1020
QY 1146 AATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATCTTGAATGATGTGTCAGG 1205
|||||
Db 1021 AATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATCTTGAATGATGTGTCAGG 1080
QY 1206 TTTGATGATGACAATCTCTCTTATCCAAATTCGCTCAGTTGCCAAGAACATCTCT 1265
|||||
Db 1081 TTTGATGATGACAATCTCTCTTATCCAAATTCGCTCAGTTGCCAAGAACATCTCT 1140
QY 1266 AAAACTTGGGTACATTTACATTGCTGCTGAAGAGAGGACTGGGACTATGCTCCCTTAGTC 1325
|||||
Db 1141 AAAACTTGGGTACATTTACATTGCTGCTGAAGAGAGGACTGGGACTATGCTCCCTTAGTC 1200
QY 1326 CTCGCCCCGATGACAGAAGTTATAAAGTCAATATTGAACAAATGGCCCTCAGCGGATT 1385
|||||
Db 1201 CTCGCCCCGATGACAGAAGTTATAAAGTCAATATTGAACAAATGGCCCTCAGCGGATT 1260
QY 1386 GGTAGGAAGTACAAAAAGTCCGATTTATGGCATACACAGATGAAAACCTTTAAGACTCGT 1445
|||||
Db 1261 GGTAGGAAGTACAAAAAGTCCGATTTATGGCATACACAGATGAAAACCTTTAAGACTCGT 1320
QY 1446 GAAGCTATTACAGATGAATCAGGAATCTTGGGACCTTACTTTATGGGGAAGTTGAGAGC 1505

Db 1321 GAAGCTATTGACGATGAATCAGGAATCTTGGACCTTACTTTATGGGGAAGTGGAGAC 1380
QY 1506 ACACGTGTGATTATATTTAAGAAATCAAGCAAGCAGACCATATACATCTACCCTCAGCGA 1565
Db 1381 ACACGTGTGATTATATTTAAGAAATCAAGCAAGCAGACCATATACATCTACCCTCAGCGA 1440
QY 1566 ATCACTGATGTCGGTCTTGTATCAAGAGATATCCAAAAGGTGTAAAACATTTGAAG 1625
Db 1441 ATCACTGATGTCGGTCTTGTATCAAGAGATATCCAAAAGGTGTAAAACATTTGAAG 1500
UY 1626 GATTTTCCAAATCTGCCCCAGAGAAATATTCAAATATTAATGACAGTGAAGTGTAGAAGAT 1685
Db 1501 GATTTTCCAAATCTGCCCCAGAGAAATATTCAAATATTAATGACAGTGAAGTGTAGAAGAT 1560
QY 1686 GGGCCAACTAAATCAGATCCTCGTGCCCTGACCCGCTATTACTCTAGTTTCTGTTAATATG 1745
Db 1561 GGGCCAACTAAATCAGATCCTCGTGCCCTGACCCGCTATTACTCTAGTTTCTGTTAATATG 1620
QY 1746 GAGAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCCCTCATCTGCTACAAAGAATCTGTA 1805
Db 1621 GAGAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCCCTCATCTGCTACAAAGAATCTGTA 1680
QY 1806 GATCAAAAGAGAAACAGATATATGTCAGACAGAGAAATGTCATCCTGTTTCTGTAATTT 1865
Db 1681 GATCAAAAGAGAAACAGATATATGTCAGACAGAGAAATGTCATCCTGTTTCTGTAATTT 1740
QY 1866 GATGAGAACCCGAAGCTGTACCTCAGAGAAATATACAGCCTTTCTCCCAATCCAGCT 1925
Db 1741 GATGAGAACCCGAAGCTGTACCTCAGAGAAATATACAGCCTTTCTCCCAATCCAGCT 1800
QY 1926 GGAGTGCAGCTTGAGGATCCAGAGTTCCAAAGCCTCCAAATCATGACAGCATCATGCGC 1985
Db 1801 GGAGTGCAGCTTGAGGATCCAGAGTTCCAAAGCCTCCAAATCATGACAGCATCATGCGC 1860
QY 1986 TATGTTTTTGTATGATTTTGCAAGTTGTCAGTTGTTGTCATGAGGTGCGATACTGCTACATT 2045
Db 1861 TATGTTTTTGTATGATTTTGCAAGTTGTCAGTTGTTGTCATGAGGTGCGATACTGCTACATT 1920
QY 2046 CTAAGCATTGGAGCAGACAGACTGACTTCTTCTGTCCTTCTCTGATATACCTTCAAA 2105
Db 1921 CTAAGCATTGGAGCAGACAGACTGACTTCTTCTGTCCTTCTCTGATATACCTTCAAA 1980
QY 2106 CACAAATGCTCTATGAAGACACACTCAACCTATTCCTCATCTCAGAGAAACTGCTTTC 2165
Db 1981 CACAAATGCTCTATGAAGACACACTCAACCTATTCCTCATCTCAGAGAAACTGCTTTC 2040
QY 2166 ATGTCGATGAAAAACCCAGGTCTATGGATTCTGGGGTGCACAACTCAGACTTTGGAAC 2225
Db 2041 ATGTCGATGAAAAACCCAGGTCTATGGATTCTGGGGTGCACAACTCAGACTTTGGAAC 2100
QY 2226 AGAGCATGACCGCTTACTGAAGTTTCTAGTTGTCAGAAAGAACTGCTGATTATAC 2285
Db 2101 AGAGCATGACCGCTTACTGAAGTTTCTAGTTGTCAGAAAGAACTGCTGATTATAC 2160
QY 2286 GAGGACAGTTATGAAGATATTTACAGCATCTGCTGAGTAAAAACAATGCCATTGAACCA 2345
Db 2161 GAGGACAGTTATGAAGATATTTACAGCATCTGCTGAGTAAAAACAATGCCATTGAACCA 2220
QY 2346 AGAAGCTTCTCCAGAAACCCAGCTCTGAAACGCCATCAAGGGAAATAACTGCTACT 2405
Db 2221 AGA-----GAATTAACCTGTAAT 2238
QY 2406 ACTCTTCAGTCAGATCAAGAGAAATGACTATGATGATACCATATCAGTTGAAATGAAG 2465
Db 2239 ACTCTTCAGTCAGATCAAGAGAAATGACTATGATGATACCATATCAGTTGAAATGAAG 2298
QY 2466 AAGGAGATTTTGAATTTATGATGAGGATGAAAAATCAGAGCCCGCAGCTTTCAAAAG 2525
Db 2299 AAGGAGATTTTGAATTTATGATGAGGATGAAAAATCAGAGCCCGCAGCTTTCAAAAG 2358
QY 2526 AAAACACGACACTATTTTATGCTGAGTGAGAGGCTCTGGATATGAGATGAGTACG 2585

Db 2359 AAAACACGACACTATTTTATGCTGACGTGAGAGGCTCTGGATATGAGATGAGTAC 2418
QY 2586 TCCCAACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTT 2645
Db 2419 TCCCAACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTT 2478
QY 2646 GTTTTCCAGGAATTTACTGATGGCTCTTACTCAGCCCTTATACCGTGGAGAACTAAAT 2705
Db 2479 GTTTTCCAGGAATTTACTGATGGCTCTTACTCAGCCCTTATACCGTGGAGAACTAAAT 2538
QY 2706 GAACATTTGGACTCTGGGCGCATATATTAAGACAGAAAGTTGAAGATAATATCATGTA 2765
Db 2539 GAACATTTGGACTCTGGGCGCATATATTAAGACAGAAAGTTGAAGATAATATCATGTA 2598
QY 2766 ACTTTCAGAAATCAGCCCTCTGTCCTTATCTCTATTTCTAGCCTTATTTCTATGAG 2825
Db 2599 ACTTTCAGAAATCAGCCCTCTGTCCTTATCTCTATTTCTAGCCTTATTTCTATGAG 2658
QY 2826 GAAGATCAGAGGCAAGGAGCAGAACTAGAAAAAACTTTGTCAAGCCTAATGAACCAAA 2885
Db 2659 GAAGATCAGAGGCAAGGAGCAGAACTAGAAAAAACTTTGTCAAGCCTAATGAACCAAA 2718
QY 2886 ACTTACTTTTGGAAAGTGCACAAATCATATGACACCCCACTAAAGATGAGTTTGAATGCAAA 2945
Db 2719 ACTTACTTTTGGAAAGTGCACAAATCATATGACACCCCACTAAAGATGAGTTTGAATGCAAA 2778
QY 2946 GCGTGGCCTTATTTCTCTGATGTTGACCTGGAAGAAAGATGTGCACTCAGGCCGTGATTGGA 3005
Db 2779 GCGTGGCCTTATTTCTCTGATGTTGACCTGGAAGAAAGATGTGCACTCAGGCCGTGATTGGA 2838
QY 3006 CCCCTTCTGCTGCGCACACTAACACACTGAACCTGCTCATGGGAGACAGTGAAGTA 3065
Db 2839 CCCCTTCTGCTGCGCACACTAACACACTGAACCTGCTCATGGGAGACAGTGAAGTA 2898
QY 3066 CAGGAATTTGCTGCTTTT--TCACCATCTTGTGATGAGACCAAAAGCTGCTACTCACT 3122
Db 2899 CAGGAATTTGCTGCTTTTTCCTCACCATCTTGTATGAGACCAAAAGCTGCTACTCACT 2958
QY 3123 GAAATATGAAAGAAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTT 3182
Db 2959 GAAATATGAAAGAAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTT 3018
QY 3183 AAAGAGAAATATGCTTCCATGCAATCAATGCTCATATGATGATACACTACCTGCTTA 3242
Db 3019 AAAGAGAAATATGCTTCCATGCAATCAATGCTCATATGATGATACACTACCTGCTTA 3078
QY 3243 GTAATGCTCAGATCAAGGATTCGATGCTATCTGCTCAGCATGGGCAGCAATGAAAC 3302
Db 3079 GTAATGCTCAGATCAAGGATTCGATGCTATCTGCTCAGCATGGGCAGCAATGAAAC 3138
QY 3303 ATCCATTTCTATTTCTAGTGACATGTGTTCACTGTACGAAAAAAAGAGAGATATAA 3362
Db 3139 ATCCATTTCTATTTCTAGTGACATGTGTTCACTGTACGAAAAAAAGAGAGATATAA 3198
QY 3363 ATGCCACTGTACATCTCTATCCAGGTGTTTTTGAGACAGTGAATGTTACCATCCAAA 3422
Db 3199 ATGCCACTGTACATCTCTATCCAGGTGTTTTTGAGACAGTGAATGTTACCATCCAAA 3258
QY 3423 GCTGGAATTTGGCGGTGGAATGCTTATTTGGCGAGCATCTACATGCTGGGATGACACA 3482
Db 3259 GCTGGAATTTGGCGGTGGAATGCTTATTTGGCGAGCATCTACATGCTGGGATGACACA 3318
QY 3483 CTTTTCCTGCTGTCAGCAATAAGTGTACAGCTCCCTGGGAATGGCTTCTGACACATT 3542
Db 3319 CTTTTCCTGCTGTCAGCAATAAGTGTACAGCTCCCTGGGAATGGCTTCTGACACATT 3378
QY 3543 AGAGATTTTTCAGATTTACAGCTTCAGGACAAATATGACAGTGGGCCCAAGCTGGCCAGA 3602
Db 3379 AGAGATTTTTCAGATTTACAGCTTCAGGACAAATATGACAGTGGGCCCAAGCTGGCCAGA 3438
QY 3603 CTTCAATATTCGGATCAATCAATGCTTGAGACCAAGAGCCCTTTCTTGATCAAG 3662
Db 3439 CTTCAATATTCGGATCAATCAATGCTTGAGACCAAGAGCCCTTTCTTGATCAAG 3498

QY 3663 GTGATCTGTGGCACCACCAATGATATTACGGCATCAAGACCCAGGGTCCCGTCAGAG 3722
|||||
Db 3499 GTGATCTGTGGCACCACCAATGATATTACGGCATCAAGACCCAGGGTCCCGTCAGAG 3558
QY 3723 TTCTCCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGAGAAGTGG 3782
|||||
Db 3559 TTCTCCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGAGAAGTGG 3618
QY 3783 CAGACTTATCGAGGAATTCACCTGGAACCTTAATGGTCTCTTTGGCAATGTGATTCA 3842
|||||
Db 3619 CAGACTTATCGAGGAATTCACCTGGAACCTTAATGGTCTCTTTGGCAATGTGATTCA 3678
QY 3843 TCTGGATAAACAATATTTTAAACCCTCAATTATGCTCGATACATCCGTTTGAC 3902
|||||
Db 3679 TCTGGATAAACAATATTTTAAACCCTCAATTATGCTCGATACATCCGTTTGAC 3738
3903 CCAACTCATATTAGCATTCGACGACTCTTCGCATGGAGTTGATGGGCTGTGATTAAAT 3962
|||||
3739 CCAACTCATATTAGCATTCGACGACTCTTCGCATGGAGTTGATGGGCTGTGATTAAAT 3798
QY 3963 AGTTCAGCATGCCATTTGGGAATGGAGAGTAAGCAATATCAGATGCACAGATTACTGCT 4022
|||||
Db 3799 AGTTCAGCATGCCATTTGGGAATGGAGAGTAAGCAATATCAGATGCACAGATTACTGCT 3858
QY 4023 TCATCCTACTTTACCAATATGTTTGGCCACTGCTCTCTTCAAAAAGCTGCACTTCACCTC 4082
|||||
Db 3859 TCATCCTACTTTACCAATATGTTTGGCCACTGCTCTCTTCAAAAAGCTGCACTTCACCTC 3918
QY 4083 CAAGGAGAGTAATGCTGGAGAGCCTGAGAGTGAATATCCAAAAGAGTGGCTGCAAGTG 4142
|||||
Db 3919 CAAGGAGAGTAATGCTGGAGAGCCTGAGAGTGAATATCCAAAAGAGTGGCTGCAAGTG 3978
QY 4143 GACTTCCAGAGACAATGAAGTCAACAGAGTAATCTCAGGAGTAATAATCTCTGCTT 4202
|||||
Db 3979 GACTTCCAGAGACAATGAAGTCAACAGAGTAATCTCAGGAGTAATAATCTCTGCTT 4038
QY 4203 ACCAGCATGTATGTGAAGAGTTCCTCATCTCCAGAGTCAAGATGGCCATCAGTGACT 4262
|||||
Db 4039 ACCAGCATGTATGTGAAGAGTTCCTCATCTCCAGAGTCAAGATGGCCATCAGTGACT 4098
QY 4263 CTCTTTTTCAGAAATGGCAAGTATTTTTCAGGGAATCAAGACTCCTTCACACCT 4322
|||||
Db 4099 CTCTTTTTCAGAAATGGCAAGTATTTTTCAGGGAATCAAGACTCCTTCACACCT 4158
QY 4323 GTGTGAACCTCTAGACCCACCGTACTGACTCGCTACCTTGAATTCACCCCCAGAGT 4382
|||||
Db 4159 GTGTGAACCTCTAGACCCACCGTACTGACTCGCTACCTTGAATTCACCCCCAGAGT 4218
QY 4383 TGGTGACACAGATGGCCCTGAGAGTGGAGTCTGGGCTGCGAGGCAAGGACCTTAC 4442
|||||
Db 4219 TGGTGACACAGATGGCCCTGAGAGTGGAGTCTGGGCTGCGAGGCAAGGACCTTAC 4278

RESULT 3
LOCUS 108643 4281 bp
DEFINITION Sequence 3 from Patent WO 8800831. PAT 14-NOV-1994
ACCESSION 108643
NID 9588649
VERSION 108643.1 GI:588649
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 4281)
AUTHORS Pasek,M.P.
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR
POLYPEPTIDES IN HIGH YIELDS

JOURNAL Patent: WO 8800831-A 3 11-FEB-1988;
FEATURES Location/Qualifiers
source 1..4281

BASE COUNT 1245 a 943 c 946 g 1147 t
ORIGIN
Query Match 86.7%; Score 4190.8; DB 5; Length 4281;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 4270; Conservative 0; Mismatches 2; Indels 48; Gaps 3;
QY 129 GCCACAGAAGATACCTGCGTGGAGTGAACATGATGAGTATATGCAAGTGAT 188
|||||
Db 4 GCCACAGAAGATACCTGCGTGGAGTGAACATGATGAGTATATGCAAGTGAT 63
QY 189 CTCGGTAGCTGCGTGGAGCGCAAGATTTCCCTAGAGTGGCAAAATCTTTCCATTTC 248
|||||
Db 64 CTCGGTAGCTGCGTGGAGCGCAAGATTTCCCTAGAGTGGCAAAATCTTTCCATTTC 123
QY 249 AACACCTCAGTGTGTACAAAAGAGCTGTGTAGAAATTCACGGATCACCTTTTCAAC 308
|||||
Db 124 AACACCTCAGTGTGTACAAAAGAGCTGTGTAGAAATTCACGGATCACCTTTTCAAC 183
QY 309 ATCGCTAAGCCAAAGCCACCCCTGGATGGGTCTGCTAGGTCTACCATCCAGGCTGAGGTT 368
|||||
Db 184 ATCGCTAAGCCAAAGCCACCCCTGGATGGGTCTGCTAGGTCTACCATCCAGGCTGAGGTT 243
QY 369 TATGATACAGTGTGTATTAACATTAAGACATGGCTTCCCATCTGCTCAGTCTTCATGCT 428
|||||
Db 244 TATGATACAGTGTGTATTAACATTAAGACATGGCTTCCCATCTGCTCAGTCTTCATGCT 303
QY 429 GTTGTGTATCTCTACGGAAGCTTCTGAGGAGCTGAATATGATGATCAGACAGTCAA 488
|||||
Db 304 GTTGTGTATCTCTACGGAAGCTTCTGAGGAGCTGAATATGATGATCAGACAGTCAA 363
QY 489 AGGGAAGAAAGATGATTAAGTCTTCCCTGTGTGAAGCCATACATATGTCGGCAGGTC 548
|||||
Db 364 AGGGAAGAAAGATGATTAAGTCTTCCCTGTGTGAAGCCATACATATGTCGGCAGGTC 423
QY 549 CTGAAAGAAATGTGCCAATGGCTCTGACCCACTGTGCTTACTACTATCTTTCT 608
|||||
Db 424 CTGAAAGAAATGTGCCAATGGCTCTGACCCACTGTGCTTACTACTATCTTTCT 483
QY 609 CATGTGACCTGTGAAAGACTGAAATTCAGGCTCATTTGAGCCCTACTAGATGTAGA 668
|||||
Db 484 CATGTGACCTGTGAAAGACTGAAATTCAGGCTCATTTGAGCCCTACTAGATGTAGA 543
QY 669 GAAAGGAGTGTGGCCAAAGAAAGACACAGACCTTGACAAATTTACTACTTTTGTCT 728
|||||
Db 544 GAAAGGAGTGTGGCCAAAGAAAGACACACAGACCTTGACAAATTTACTACTTTTGTCT 603
QY 729 GTAATTTGATGAAGGAAAGTGGCACTCAGAAACAAAGAACTCCTTGATGAGGATAGG 788
|||||
Db 604 GTAATTTGATGAAGGAAAGTGGCACTCAGAAACAAAGAACTCCTTGATGAGGATAGG 663
QY 789 GATGCTGATCTGCTGGGCTGGCTTAAATGCACACAGTCAATGTTATGTAACAGG 848
|||||
Db 664 GATGCTGATCTGCTGGGCTGGCTTAAATGCACACAGTCAATGTTATGTAACAGG 723
QY 849 TCTCTG---CCAGCTGTGATGGATGGCAGCAAGAAATCAGTCTATTGGCATGTGATGGA 905
|||||
Db 724 TCTCTGCTACCAAGGCTGTGATGGATGGCAGCAAGAAATCAGTCTATTGGCATGTGATGGA 783
QY 906 ATGGGACACACCTCTGAAGTGCATCAATATTCCTGGAAGTGCACACATTTCTTGAGG 965
|||||
Db 784 ATGGGACACACCTCTGAAGTGCATCAATATTCCTGGAAGTGCACACATTTCTTGAGG 843
QY 966 AACCATGCGCAGGCGCTTGGAAATCTGCGCAATAACTTCTTACTGCTCAACACTC 1025
|||||
Db 844 AACCATGCGCAGGCGCTTGGAAATCTGCGCAATAACTTCTTACTGCTCAACACTC 903
QY 1026 TTGATGACCTTGGACAGTTTCTACTGTTTGTATATATCTTCCACCAACATGATGGC 1085
|||||
Db 904 TTGATGACCTTGGACAGTTTCTACTGTTTGTATATATCTTCCACCAACATGATGGC 963

QY 1086 ATGGAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGGAACCCCACTACGAATGAAAAAT 1145
|||||
Db 964 ATGGAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGGAACCCCACTACGAATGAAAAAT 1023
QY 1146 AATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAAATGATGTGTGACAG 1205
|||||
Db 1024 AATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAAATGATGTGTGACAG 1083
QY 1206 TTTGATGATGACAACCTCTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCT 1265
|||||
Db 1084 TTTGATGATGACAACCTCTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCT 1143
QY 1266 AAAACTTGGGTACATTACATTTGCTGTGAAGAGAGAGACTGGACTATGCTCCCTTAAGTC 1325
|||||
Db 1144 AAAACTTGGGTACATTACATTTGCTGTGAAGAGAGAGACTGGACTATGCTCCCTTAAGTC 1203
QY 1326 CTCGCCCCGATGACAGAAGTTATATAAGTCAATATTTGAACAATGGCCCTCAGCGGAT 1385
|||||
Db 1204 CTCGCCCCGATGACAGAAGTTATATAAGTCAATATTTGAACAATGGCCCTCAGCGGAT 1263
QY 1386 GGTAGAGTACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGT 1445
|||||
Db 1264 GGTAGAGTACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGT 1323
QY 1446 GAAGCTATTACAGATGATCAGGAATCTGGGACCTTTACTTTATGGGAAAGTTGAGAC 1505
|||||
Db 1324 GAAGCTATTACAGATGATCAGGAATCTGGGACCTTTACTTTATGGGAAAGTTGAGAC 1383
QY 1506 ACACGTGTGATTAATTAAGATCAAGCAAGCAGACCATATAACATCTACCCCTCAGCGA 1565
|||||
Db 1384 ACACGTGTGATTAATTAAGATCAAGCAAGCAGACCATATAACATCTACCCCTCAGCGA 1443
QY 1566 ATCAGTATGTCCGCTTTGTATTCAGAGAGATTACCAAAAGGTGTAACAATTTGAAG 1625
|||||
Db 1444 ATCAGTATGTCCGCTTTGTATTCAGAGAGATTACCAAAAGGTGTAACAATTTGAAG 1503
QY 1626 GATTTTCCAATTCTGCCAGAGAAATATTCAATATAAATGACAGTACTGTAGAGAT 1685
|||||
Db 1504 GATTTTCCAATTCTGCCAGAGAAATATTCAATATAAATGACAGTACTGTAGAGAT 1563
QY 1686 GGGCAACTAAATCAGATCTCGGTGCTGACCCGCTATTACTCTAGTTTCTTAATATG 1745
|||||
Db 1564 GGGCAACTAAATCAGATCTCGGTGCTGACCCGCTATTACTCTAGTTTCTTAATATG 1623
QY 1746 GAGAGAGATCTAGCTTCAGAGACTCATTTGGCCCTCTCTCATCTGCTACAAGAATCTGTA 1805
|||||
Db 1624 GAGAGAGATCTAGCTTCAGAGACTCATTTGGCCCTCTCTCATCTGCTACAAGAATCTGTA 1683
QY 1805 GATCAAGAAGAAACAGATATATGTCAACAAGAGAATGTCATCTCTTTTCTGTATTT 1865
|||||
Db 1684 GATCAAGAAGAAACAGATATATGTCAACAAGAGAATGTCATCTCTTTTCTGTATTT 1743
QY 1866 GATGAGAACCGAAGCTGTACTCTACAGAGAATATACAAGCTTTCTCCCAATCCAGCT 1925
|||||
Db 1744 GATGAGAACCGAAGCTGTACTCTACAGAGAATATACAAGCTTTCTCCCAATCCAGCT 1803
QY 1926 GGAGTGCAGCTTGAGATCCAGAGTTCCAAAGCCTCCAACATCATGACAGCATCAATGGC 1985
|||||
Db 1804 GGAGTGCAGCTTGAGATCCAGAGTTCCAAAGCCTCCAACATCATGACAGCATCAATGGC 1863
QY 1986 TATGTTTTGATAGTTTGCAGTTGTGTGAGTTTGTGCAATGAGGTGGCATACTGGTACATT 2045
|||||
Db 1864 TATGTTTTGATAGTTTGCAGTTGTGTGAGTTTGTGCAATGAGGTGGCATACTGGTACATT 1923
QY 2046 CTAAGCATTTGGAGCACAGACTGACTTCTTCTGTCTTCTCTGTGATATACCTTCAA 2105
|||||
Db 1924 CTAAGCATTTGGAGCACAGACTGACTTCTTCTGTCTTCTCTGTGATATACCTTCAA 1983
QY 2106 CACAAATGCTTATGAGACACACTCACTCCCTATTTCCCATTTCTCAGGAGAAAACTGTCTTC 2165
|||||
Db 1984 CACAAATGCTTATGAGACACACTCACTCCCTATTTCCCATTTCTCAGGAGAAAACTGTCTTC 2043
QY 2166 ATGTCATGAAAAACCCAGGCTTATGATTTCTGGGTGCCACAACCTCAGACTTTGGGAAC 2225

Db 2044 ATGTCATGAAAAACCCAGGCTATGATTTCTGGGTGCCACACACTCAGACTTTCCGAAC 2103
QY 2226 AGAGGATGACCCGCTTACTGAAGGTTTCTAGTTGTGACAAGAACACTGGTATTTAC 2285
|||||
Db 2104 AGAGGATGACCCGCTTACTGAAGGTTTCTAGTTGTGACAAGAACACTGGTATTTAC 2163
QY 2286 GAGGACGTTATGAGATATTTACAGTACTTGTCTGAGTAAAAACAATGCCATTGAAACA 2345
|||||
Db 2164 GAGGACGTTATGAGATATTTACAGTACTTGTCTGAGTAAAAACAATGCCATTGAAACA 2223
QY 2346 AGAAGCTTCTCCAGAACCCACAGTCTTGAACGCCATCAACGGGAATTAAGTCTGACT 2405
|||
Db 2224 AGA-----GAAATACTGACT 2241
QY 2406 ACTCTTCAGTCAGATCAAGAGAAATTTGACTATGATATACCATATCAGTTGAATGAAG 2465
|||||
Db 2242 ACTCTTCAGTCAGATCAAGAGAAATTTGACTATGATATACCATATCAGTTGAATGAAG 2301
QY 2466 AAGGACATTTGACATTTATGATGAGGATGAAAAATCAGAGCCCCCGCAGCTTCAAAAG 2525
|||||
Db 2302 AAGGACATTTGACATTTATGATGAGGATGAAAAATCAGAGCCCCCGCAGCTTCAAAAG 2361
QY 2526 AAAACACGACACTATTTTATTTGCTGCAAGTGAGAGGCTCTGGGATTAATGGATGAGTAGC 2585
|||||
Db 2362 AAAACACGACACTATTTTATTTGCTGCAAGTGAGAGGCTCTGGGATTAATGGATGAGTAGC 2421
QY 2586 TCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTT 2645
|||||
Db 2422 TCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTT 2481
QY 2646 GTTTTCCAGGAATTTACTGATGCTCTTACTCTACAGCCCTTATACCGTGGAGAACTAAAT 2705
|||||
Db 2482 GTTTTCCAGGAATTTACTGATGCTCTTACTCTACAGCCCTTATACCGTGGAGAACTAAAT 2541
QY 2706 GAACATTTGGACTCTGGGCCATATTAAGAGCAGAAAGTTGAAGATTAATATCATGTA 2765
|||||
Db 2542 GAACATTTGGACTCTGGGCCATATTAAGAGCAGAAAGTTGAAGATTAATATCATGTA 2601
QY 2766 ACTTTCAGAAATCAGGCTCTGTCCTTATCTCTATTTCTAGCCTTATTTCTTATGAG 2825
|||||
Db 2602 ACTTTCAGAAATCAGGCTCTGTCCTTATCTCTATTTCTAGCCTTATTTCTTATGAG 2661
QY 2826 GAAGATCAGAGGCAAGGAGCAGAACCTAGAAAAAATTGTCAAGCCTAATGAACCAAA 2885
|||||
Db 2662 GAAGATCAGAGGCAAGGAGCAGAACCTAGAAAAAATTGTCAAGCCTAATGAACCAAA 2721
QY 2886 ACTTACTTTTGAAGGTGCAACATCATATGACACCACTAAAGATGAGTTTGACTGCAAA 2945
|||||
Db 2722 ACTTACTTTTGAAGGTGCAACATCATATGACACCACTAAAGATGAGTTTGACTGCAAA 2781
QY 2946 GCCTGGCTTATTTCTGTGATGTTGAGTGAAGAAAGATGTGCACACTGAGGCTGATGGA 3005
|||||
Db 2782 GCCTGGCTTATTTCTGTGATGTTGAGTGAAGAAAGATGTGCACACTGAGGCTGATGGA 2841
QY 3006 CCCCTCTGCTGTCACACTAACACACTGAACCTGCTCATGGGAGACAAGTGACAGTA 3065
|||||
Db 2842 CCCCTCTGCTGTCACACTAACACACTGAACCTGCTCATGGGAGACAAGTGACAGTA 2901
QY 3066 CAGGAATTTGCTGTGTTT--TCACCATCTTTGATGAGACCAAAAGCTGTACTTCACT 3122
|||||
Db 2902 CAGGAATTTGCTGTGTTTCTCTCACCATCTTTGATGAGACCAAAAGCTGTACTTCACT 2961
QY 3123 GAAATATGGAAGAAGAACTGCAAGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTT 3182
|||||
Db 2962 GAAATATGGAAGAAGAACTGCAAGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTT 3021
QY 3183 AAAGACATTTATCGTTCATGCAATCAATGCTACATAATGATACACTACCTGGCTTA 3242
|||||
Db 3022 AAAGACATTTATCGTTCATGCAATCAATGCTACATAATGATACACTACCTGGCTTA 3081
QY 3243 GTAATGCTCAGATCAAAAGATTCGATGTTCTGCTCAGCATGGGCAGCAATGAANAAC 3302
|||||

Db 3082 GTAATGGCTCAGGATCAAAAGATTGATGTAATCTGCTCAGCATGGGCAATGAAAC 3141
QY 3303 ATCCATTCTATTCATTTCAGTGGACATGTGTCACTGTACGAAAAAAGAGAGTATAA 3362
Db 3142 ATCCATTCTATTCATTTCAGTGGACATGTGTCACTGTACGAAAAAAGAGAGTATAA 3201
QY 3363 ATGGCACTGTACAAATCTCTATCCAGGTGTTTGTGAGACAGTGAATGTTACATCCAA 3422
Db 3202 ATGGCACTGTACAAATCTCTATCCAGGTGTTTGTGAGACAGTGAATGTTACATCCAA 3261
QY 3423 GCTGAATTTGGCGGGTGGAAATGCTTATTTGGCGAGCATCTACATGCTGGGATGAGCACA 3482
Db 3262 GCTGAATTTGGCGGGTGGAAATGCTTATTTGGCGAGCATCTACATGCTGGGATGAGCACA 3321
QY 3483 CTTTCTGTGTGTACAGCAATAAGTGTACAGACTCCCTGGGAATGGCTTCTGACACATT 3542
Db 3322 CTTTCTGTGTGTGTACAGCAATAAGTGTACAGACTCCCTGGGAATGGCTTCTGACACATT 3381
QY 3543 AGAGATTTTGTAGATTACAGCTTCAGGACAATATGGACAGTGGGCCCCAAAGCTGGCCAGA 3602
Db 3382 AGAGATTTTGTAGATTACAGCTTCAGGACAATATGGACAGTGGGCCCCAAAGCTGGCCAGA 3441
QY 3603 CTTCAATTATCCGGATCAATCAATGCTGTGAGCACCAGAGCCCTTTCTTGTGATCAAG 3662
Db 3442 CTTCAATTATCCGGATCAATCAATGCTGTGAGCACCAGAGCCCTTTCTTGTGATCAAG 3501
QY 3663 GTGGATCTGTGGCACCAGATGATTTATTCAGGCGCATCAAGACCAGGGGTGCCGTGAGAG 3722
Db 3502 GTGGATCTGTGGCACCAGATGATTTATTCAGGCGCATCAAGACCAGGGGTGCCGTGAGAG 3561
QY 3723 TTCTCCAGCCTCTACATCTCTCAGTTTATATCATGTATAGTCTTGATGGGAGAAGTGG 3782
Db 3562 TTCTCCAGCCTCTACATCTCTCAGTTTATATCATGTATAGTCTTGATGGGAGAAGTGG 3621
QY 3783 CAGACTTATCGAGGAATAATTCACCTGGAACCTTAATGGTCTTCTTGGCAATGTGATTCA 3842
Db 3622 CAGACTTATCGAGGAATAATTCACCTGGAACCTTAATGGTCTTCTTGGCAATGTGATTCA 3681
QY 3843 TCTGGGATATAACACAATATTTTAAACCTTCCAATTATTTGCTCGATACATCCGTTTGAC 3902
Db 3682 TCTGGGATATAACACAATATTTTAAACCTTCCAATTATTTGCTCGATACATCCGTTTGAC 3741
QY 3903 CCAACTCATTTATAGCATTCGACGACCTCTTGCATGGAGTTGATGGGCTGTGATTTAAAT 3962
Db 3742 CCAACTCATTTATAGCATTCGACGACCTCTTGCATGGAGTTGATGGGCTGTGATTTAAAT 3801
QY 3963 AGTTCAGCATGCCATTGGGAATGGAGAGTAAAGCAATATCAGATGCACAGATTACTGCT 4022
Db 3802 AGTTCAGCATGCCATTGGGAATGGAGAGTAAAGCAATATCAGATGCACAGATTACTGCT 3861
QY 4023 TCATCTCTACTTTACCAATATGTTTGGCACCTGTGCTCTCTCAAAAGCTGCAGTTTCACTTC 4082
Db 3862 TCATCTCTACTTTACCAATATGTTTGGCACCTGTGCTCTCTCAAAAGCTGCAGTTTCACTTC 3921
QY 4083 CAAGGAGAGTAAATGCTGTGGAGACCTTCAGGTGAATAATCCAAAAAGAGTGGCTGCAAGTG 4142
Db 3922 CAAGGAGAGTAAATGCTGTGGAGACCTTCAGGTGAATAATCCAAAAAGAGTGGCTGCAAGTG 3981
QY 4143 GACTTCAGAGACAATGAAGTCAACAGAGTAACTACTCAGGAGTAAATCTCTGCTT 4202
Db 3982 GACTTCAGAGACAATGAAGTCAACAGAGTAACTACTCAGGAGTAAATCTCTGCTT 4041
QY 4203 ACCAGCATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACT 4262
Db 4042 ACCAGCATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACT 4101
QY 4263 CTCTTTTTCAGAAATGGCAAAAGTAAAGTTTTTCAGGGAATCAAGACTCTCTTACACCT 4322
Db 4102 CTCTTTTTCAGAAATGGCAAAAGTAAAGTTTTTCAGGGAATCAAGACTCTCTTACACCT 4161
QY 4323 GTGTGAACTCTCTAGACCACCGTTACTGACTGCTACTCTTGAATTCACCCCCAGAGT 4382
Db 4162 GTGTGAACTCTCTAGACCACCGTTACTGACTGCTACTCTTGAATTCACCCCCAGAGT 4221

QY 4383 TGGGTGCACCAAGATTGCCCTGAGGATGAGGTTCTGTGGCTGCGAGGACAGGACCTCTAC 4442
Db 4222 TGGGTGCACCAAGATTGCCCTGAGGATGAGGTTCTGTGGCTGCGAGGACAGGACCTCTAC 4281

RESULT 4
LOCUS 108641 4551 bp
DEFINITION Sequence 1 from Patent WO 8800831. PAT 14-NOV-1994
ACCESSION 108641
NID 9588647
VERSION 108641.1 GI:588647
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4551)
AUTHORS Pask, M.P.
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR VIII:C-LIKE POLYPEPTIDES AND PROCESSES FOR PRODUCING THESE
JOURNAL POLYPEPTIDES IN HIGH YIELDS
FEATURES Patent: WO 8800831-A 1 11-FEB-1988;
source Location/Qualifiers
1.4551 /organism="unknown"
BASE COUNT 1345 a 1003 c 1004 g 1198 t 1 others
ORIGIN

Query Match 83.7%; Score 4045.8; DB 5; Length 4551;
Best Local Similarity 94.8%; Pred. No. 0;
Matches 4311; Conservative 0; Mismatches 3; Indels 234; Gaps 3;

QY 129 GCCACGAGAAGATACACTGCTGGTGCAGTGAACCTGCATGGAGTATATGCAAAAGTAT 188
Db 4 GCCACGAGAAGATACACTGCTGGTGCAGTGAACCTGCATGGAGTATATGCAAAAGTAT 63
QY 189 CTCGTGAGCTGCCCTGTGGACGCAAGATTCTCTCTAGAGTGCACAAATCTTTTCCATT 248
Db 64 CTCGTGAGCTGCCCTGTGGACGCAAGATTCTCTCTAGAGTGCACAAATCTTTTCCATT 123
QY 249 AACACCTCAGTCTGTACAAAAGAGCTCTGTTGTAGAAATTCACGGATCACCTTTTCAAC 308
Db 124 AACACCTCAGTCTGTACAAAAGAGCTCTGTTGTAGAAATTCACGGATCACCTTTTCAAC 183
QY 309 ATCGTAAGCCAGAGGCCACCTCGATGGGTCTGCTAGGCTCTACCATCCAGGCTAGGTT 368
Db 184 ATCGTAAGCCAGAGGCCACCTCGATGGGTCTGCTAGGCTCTACCATCCAGGCTAGGTT 243
QY 369 TATGATACAGTGTGCTATACACTTAAGACATGGCTTCCCATCCCTGCTCAGTCTTATGCT 428
Db 244 TATGATACAGTGTGCTATACACTTAAGACATGGCTTCCCATCCCTGCTCAGTCTTATGCT 303
QY 429 GTTGTGTATCTCTACTGGAAGCTTCTGAGGAGCTGAATATGATCATCAGACCACTCAA 488
Db 304 GTTGTGTATCTCTACTGGAAGCTTCTGAGGAGCTGAATATGATCATCAGACCACTCAA 363
QY 489 AGGGAAGAAAGAGATGATAAGTCTTCCCTGTGTGAAGCCATACATATGTCTGGAGGTC 548
Db 364 AGGGAAGAAAGAGATGATAAGTCTTCCCTGTGTGAAGCCATACATATGTCTGGAGGTC 423
QY 549 CTGAAGAAGATGTGCCAATGTGCTCTGACCCACTGTGCTTACTACTCATATCTTTCT 608
Db 424 CTGAAGAAGATGTGCCAATGTGCTCTGACCCACTGTGCTTACTACTCATATCTTTCT 483
QY 609 CATGTGACCTGTGTAAGAAGACTTGAATTCAGGCTCATTTGAGGCCCTACTAGTATGTAGA 668
Db 484 CATGTGACCTGTGTAAGAAGACTTGAATTCAGGCTCATTTGAGGCCCTACTAGTATGTAGA 543
QY 669 GAAGGAGTCTGGCCAAAGAAAGACACAGACCTTGCACAAATTTATACTACTTTTGTCT 728
Db 544 GAAGGAGTCTGGCCAAAGAAAGACACACACCTTTGCACAAATTTATACTACTTTTGTCT 603

QY	729	GTATTGATGAAGGAAAAAGTTGGCACTCAGAAACAAAGAACTCCTTGATGCAGATAGG	788
Db	604	GTATTGATGAAGGAAAAAGTTGGCACTCAGAAACAAAGAACTCCTTGATGCAGATAGG	663
QY	789	GATGTCGATCTGCTCGGGCCCTGGCCTAAATGCAACAGTCAGTCAATGGTTATGTAACAGG	848
Db	664	GATGTCGATCTGCTCGGGCCCTGGCCTAAATGCAACAGTCAGTCAATGGTTATGTAACAGG	723
QY	849	TCTCTG--CCAGTCTGATTGGATGCCCACAGAGAAATCAGTCTAFTGGCATGTGATTGGA	905
Db	724	TCTCTGCTACCAAGTCTGATTGGATGCCCACAGAGAAATCAGTCTAFTGGCATGTGATTGGA	783
QY	906	ATGGCACCACCTCCTGAAGTGCACATCATATTTCTCGAAGGTCAACATTTCTGTGAGG	965
Db	784	ATGGCACCACCTCCTGAAGTGCACATCATATTTCTCGAAGGTCAACATTTCTGTGAGG	843
QY	966	AACCATCGCCAGGCGTCTTGAATCTCGCAATACTTCTTACTGTCTCAAACTC	1025
Db	844	AACCATCGCCAGGCGTCTTGAATCTCGCAATACTTCTTACTGTCTCAAACTC	903
Db	1026	TTGATGGACCTTGGACAGTTTCTACTGTTTGTCAATCTCTTCCCAACACATGATGGC	1089
QY	904	TTGATGGACCTTGGACAGTTTCTACTGTTTGTCAATCTCTTCCCAACACATGATGGC	963
QY	1086	ATGGAAGCTTATGTCAAGTAGACAGCTGTCCAGAGGAACCCCACTACGAATGAAAAAT	1145
Db	964	ATGGAAGCTTATGTCAAGTAGACAGCTGTCCAGAGGAACCCCACTACGAATGAAAAAT	1023
QY	1146	AATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAATGGATGTGCTCAGG	1209
Db	1024	AATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAATGGATGTGCTCAGG	1083
QY	1206	TTTGTGATGACAACTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGACATCCT	1265
Db	1084	TTTGTGATGACAACTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGACATCCT	1143
QY	1266	AAAACCTGGGTACATTTACATTGCTGCTGAAGAGAGAGACTGGACTATGCTCCCTTAGTC	1329
Db	1144	AAAACCTGGGTACATTTACATTGCTGCTGAAGAGAGAGACTGGACTATGCTCCCTTAGTC	1203
QY	1326	CTCGCCCCCGATGACAGAAAGTTATAAAGTCAATATTGACAAATGGCCCTCAGCGGATT	1389
Db	1204	CTCGCCCCCGATGACAGAAAGTTATAAAGTCAATATTGACAAATGGCCCTCAGCGGATT	1265
QY	1386	GGTAGGAAGTACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGT	1445
Db	1264	GGTAGGAAGTACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGT	1323
QY	1446	GAACTATTACGATGAATCAGAAATCTTGGGACCTTACTTTATGGGGAAGTTGGAGAC	1509
Db	1324	GAACTATTACGATGAATCAGAAATCTTGGGACCTTACTTTATGGGGAAGTTGGAGAC	1383
QY	1506	ACACTGTTGATTATATTAAGAATCAAGCAAGCAGACCATATTAACATCTACCCCTCACGGA	1565
Db	1384	ACACTGTTGATTATATTAAGAATCAAGCAAGCAGACCATATTAACATCTACCCCTCACGGA	1443
QY	1566	ATCACTGATGTCGGTCTTTGTATTCAGGAGATTAACCAAAAGGTGTAACATTTGAAG	1622
Db	1444	ATCACTGATGTCGGTCTTTGTATTCAGGAGATTAACCAAAAGGTGTAACATTTGAAG	1503
QY	1626	GATTTTCCAATTCTGCGCAGGAGAAATATTCMAAATAAATGAGACGTGATGAGAAAT	1689
Db	1504	GATTTTCCAATTCTGCGCAGGAGAAATATTCMAAATAAATGAGACGTGATGAGAAAT	1563
QY	1686	GGGCCAACTAAATCAGATCCTCGGTGCTGACCCGCTATTACTCTAGTTTCGTTAATATG	1745
Db	1564	GGGCCAACTAAATCAGATCCTCGGTGCTGACCCGCTATTACTCTAGTTTCGTTAATATG	1623
QY	1746	GAGAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGCTACAAAGAACTGTGA	1805
Db	1624	GAGAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGCTACAAAGAACTGTGA	1683

QY	1806	GATCAAGAGGAAACCCAGATAATGTCTGACACAGAGGAAGATGTCATCTGTCTTCTGTATT	1865
Db	1684	GATCAAGAGGAAACCCAGATAATGTCTGACACAGAGGAAGATGTCATCTGTCTTCTGTATT	1743
QY	1866	GATGAGAACCAGAACTGTGTACCTCACAGAGAAATATACAACGCTTCTCTCCCAATCCAGCT	1925
Db	1744	GATGAGAACCAGAACTGTGTACCTCACAGAGAAATATACAACGCTTCTCTCCCAATCCAGCT	1803
QY	1926	GGAGTGCAGCTTGAGGATCCAGAGTCCAAAGCCTTCCAACATCATGCAACAGCATCAATGCG	1985
Db	1804	GGAGTGCAGCTTGAGGATCCAGAGTCCAAAGCCTTCCAACATCATGCAACAGCATCAATGCG	1863
QY	1986	TATGTTTTGTAGTATTGTGCAGTGTGTCTCAGTTTGTTCATGAGGTGGCATACTGTACATT	2045
Db	1864	TATGTTTTGTAGTATTGTGCAGTGTGTCTCAGTTTGTTCATGAGGTGGCATACTGTACATT	1923
QY	2046	CTAAGCATTTGAGACACAGACTGACTCTCTTCTGTCTTCTCTGTGATATACCTTCAA	2105
Db	1924	CTAAGCATTTGAGACACAGACTGACTCTCTTCTGTCTTCTCTGTGATATACCTTCAA	1983
QY	2106	CACAAATGCTCTATGAAGACACACTCACCCCTATTCCTCAGAGAACTGTCTTC	2165
Db	1984	CACAAATGCTCTATGAAGACACACTCACCCCTATTCCTCAGAGAACTGTCTTC	2043
QY	2166	ATGTGATGGAACCCAGGCTATGTGATTCTGTGGGTGCCACAACCTCAGACTTTCGGAAC	2225
Db	2044	ATGTGATGGAACCCAGGCTATGTGATTCTGTGGGTGCCACAACCTCAGACTTTCGGAAC	2103
QY	2226	AGAGCATGACCGCTTACTGAAGTTCCTAGTTGTGACAAACAACACTGCTGATTATTAC	2285
Db	2104	AGAGCATGACCGCTTACTGAAGTTCCTAGTTGTGACAAACAACACTGCTGATTATTAC	2163
QY	2286	GAGCAGCTATGAAGATATTTCAGCATACTTGTGAGTAAACCAATGCCATTGAACCA	2345
Db	2164	GAGCAGCTATGAAGATATTTCAGCATACTTGTGAGTAAACCAATGCCATTGAACCA	2223
QY	2346	AGAAGCTTCTCCACG-----	2361
Db	2224	AGAAGCTTCTCCACGAGATCCTCTGTGCTTGGATTAACCACTATGTACTCAGATACCAAAA	2283
QY	2361	-----	2361
Db	2284	GAAAGTGGAAATCCCAAGAGAGTCAACGAAAAAAGCTTTTAAGAAAAAGGATACC	2343
QY	2361	-----	2361
Db	2344	ATTTTGTCCCTGAACGCTTGTGAAGCAATCATGCAATAGCAGCAATAAATGAGGACAA	2403
QY	2361	-----	2361
Db	2404	AATAAGCCCGAAATAGAAAGTCAACCTGGGCAAAAGCAAGTAGAGCTGAAAGGCTGTCTCT	2463
QY	2361	---AACCACACAGTCTTGAACGCGCATCAACGGGAAATAACTGTACTACTCTTCAGTCA	2417
Db	2464	CAAAACCCACACAGTCTTGAACGCGCATCAACGGGAAATAACTGTACTACTCTTCAGTCA	2523
QY	2418	GATCAGAGGAAATGACTATGATGATACCATATCAGTTGAAATGAAGAAGAGATTCT	2477
Db	2524	GATCAGAGGAAATGACTATGATGATACCATATCAGTTGAAATGAAGAAGAGATTCT	2583
QY	2478	GACATTTATGATGAGATGAAAAATCAGAGCCCGCAGCTTCAAAAAGAAAAACGACAC	2537
Db	2584	GACATTTATGATGAGATGAAAAATCAGAGCCCGCAGCTTCAAAAAGAAAAACGACAC	2643
QY	2538	TATTTTATTTGCTGAGTGGAGAGGCTCTGGGATTATGGGATAGTAGTCCCAATGTT	2597
Db	2644	TATTTTATTTGCTGAGTGGAGAGGCTCTGGGATTATGGGATAGTAGTCCCAATGTT	2703
QY	2598	CTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGAA	2657
Db	2704	CTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGAA	2763
QY	2658	TTTACTGATGGCTCTTACTTCAGCCCTTATACCTGTGAGAACTAAATGAACATTTGGGA	2717


```
|||||
Db 2764 TTACTGATGGCTCCTTACTCAGCCCTTATACCGTGAGAACTAAATGACATTTGGGA 2823
QY 2718 CTCCTGGGGCCATATATAGAGCAGAGAGTTGAAGATATATCATGTGTAACCTTCAGAAAT 2777
Db 2824 CTCCTGGGGCCATATATAGAGCAGAGAGTTGAAGATATATCATGTGTAACCTTCAGAAAT 2883
QY 2778 CAGGCTCTGCTCCCTATTTCTTCTATCTAGCCCTTATTTCTTATGAGAGATCAGAGG 2837
Db 2884 CAGGCTCTGCTCCCTATTTCTTCTATCTAGCCCTTATTTCTTATGAGAGATCAGAGG 2943
QY 2838 CAAGAGCAGAGAACCTAGAAAAAAGTTTGTCAAGCCCTAATGAACCAAACTTACTTTGG 2897
Db 2944 CAAGAGCAGAGAACCTAGAAAAAAGTTTGTCAAGCCCTAATGAACCAAACTTACTTTGG 3003
QY 2898 AAAGTGACATCATATATGACACCCACTAAAGATGAGTTGACTGCAAAAGCCCTGGCTTAT 2957
Db 3004 AAAGTGACATCATATATGACACCCACTAAAGATGAGTTGACTGCAAAAGCCCTGGCTTAT 3063
QY 2958 TTCTCTGATGTGACCTGGAAAAAGATGTGCACTCAGGCCCTGATTGAGCCCTTCTGCTC 3017
Db 3064 TTCTCTGATGTGACCTGGAAAAAGATGTGCACTCAGGCCCTGATTGAGCCCTTCTGCTC 3123
QY 3018 TGCCACACTAACACACTGAACCCCTGCTCATGGAGACAGTACAGTACAGGAATTTGCT 3077
Db 3124 TGCCACACTAACACACTGAACCCCTGCTCATGGAGACAGTACAGTACAGGAATTTGCT 3183
QY 3078 CTGTTT---TCACCATCTTTGATGAGACCAAAAGCTGGTACTTCACTGAAAAATATGGAA 3134
Db 3184 CTGTTTCTCTCACCACCTTTGATGAGACCAAAAGCTGGTACTTCACTGAAAAATATGGAA 3243
QY 3135 AGAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAAATAT 3194
Db 3244 AGAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAAATAT 3303
QY 3195 CGCTTCCATGCAATCAATGGCTACATATGGATACACTACCTGGCTTAGTAATGGCTCAG 3254
Db 3304 CGCTTCCATGCAATCAATGGCTACATATGGATACACTACCTGGCTTAGTAATGGCTCAG 3363
QY 3255 GATCAAGAGATTCGATGCTATCTGCTCAGCATGGGAGCAATGAACATCCATCTTAT 3314
Db 3364 GATCAAGAGATTCGATGCTATCTGCTCAGCATGGGAGCAATGAACATCCATCTTAT 3423
QY 3315 CATTCAGTGGACATGTGTTCACTGTACGAAAAAAGAGAGTATAAATGGCACTGTAC 3374
Db 3424 CATTCAGTGGACATGTGTTCACTGTACGAAAAAAGAGAGTATAAATGGCACTGTAC 3483
QY 3375 AATCTCTATCCAGTGTGTTTGGAGACAGTGGAAATGTTACCATCCAAAGCTGGAATTTGG 3434
Db 3484 AATCTCTATCCAGTGTGTTTGGAGACAGTGGAAATGTTACCATCCAAAGCTGGAATTTGG 3543
QY 3435 CGGTTGGAATGCTTATTTGGGAGACATCTACATGCTGGGATGAGCACACTTTTCTGCTG 3494
Db 3544 CGGTTGGAATGCTTATTTGGGAGACATCTACATGCTGGGATGAGCACACTTTTCTGCTG 3603
QY 3495 TACAGCAATAAGTGTACAGACTCCCTGGGAATGGCTTCTGGACACATTAGAGATTTTCAG 3554
Db 3604 TACAGCAATAAGTGTACAGACTCCCTGGGAATGGCTTCTGGACACATTAGAGATTTTCAG 3663
QY 3555 ATTACAGCTTCAGAGACAATATGAGCAGTGGGCCCCAAAAGCTGGCCAGACTTCATTATTCC 3614
Db 3664 ATTACAGCTTCAGAGACAATATGAGCAGTGGGCCCCAAAAGCTGGCCAGACTTCATTATTCC 3723
QY 3615 GGATCAATCAATGCTTGAGACACCAAGAGGCCCTTTTCTTGATCAAGGTGATCTGTTG 3674
Db 3724 GGATCAATCAATGCTTGAGACACCAAGAGGCCCTTTTCTTGATCAAGGTGATCTGTTG 3783
QY 3675 GCACCAATGATTTATTCACGGCAATCAAGACCCAGGGTGCCCGTGAGAAAGTTCTCCAGCCTC 3734
Db 3784 GCACCAATGATTTATTCACGGCAATCAAGACCCAGGGTGCCCGTGAGAAAGTTCTCCAGCCTC 3843
QY 3735 TACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGAGTGCGAGACTTATCGA 3794
|||||
```

```
Db 3844 TACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGAGTGCGAGACTTATCGA 3903
QY 3795 GGAATTTCCACTGGAAACCTTAATAGTCTTCTTTGGCAATGTGATTCATCTGGGATAAA 3854
Db 3904 GGAATTTCCACTGGAAACCTTAATAGTCTTCTTTGGCAATGTGATTCATCTGGGATAAA 3963
QY 3855 CACATATTTTAAACCTCCCAATTAATGTCTGCATACATCCGTTTGCACCAACTCATTAT 3914
Db 3964 CACATATTTTAAACCTCCCAATTAATGTCTGCATACATCCGTTTGCACCAACTCATTAT 4023
QY 3915 AGCATTCGACAGCACTCTTCCATGAGAGTGTGATGGGCTGTGATTTAAATAGTTGCAGCATG 3974
Db 4024 AGCATTCGACAGCACTCTTCCATGAGAGTGTGATGGGCTGTGATTTAAATAGTTGCAGCATG 4083
QY 3975 CCATTTGGGAATGAGAGTAAAGCAATATCAGATGCACAGATTACTGCTTCATCTTACTTT 4034
Db 4084 CCATTTGGGAATGAGAGTAAAGCAATATCAGATGCACAGATTACTGCTTCATCTTACTTT 4143
QY 4035 ACCAATATGTTGGCCACCTGGTCTCTCTCAAAAGCTGACCTTCACTCCAAAGGAGAGT 4094
Db 4144 ACCAATATGTTGGCCACCTGGTCTCTCTCAAAAGCTGACCTTCACTCCAAAGGAGAGT 4203
QY 4095 AATGCTTGAGACCTCAGTGTATATATCCAAAGAGTGGCTGCAAGTGAGCTTCCAGAA 4154
Db 4204 AATGCTTGAGACCTCAGTGTATATATCCAAAGAGTGGCTGCAAGTGAGCTTCCAGAA 4263
QY 4155 ACAATGAAAGTACAGAGAGTAACTACTCAGGAGAGTAAATCTGCTTACAGCATGTAT 4214
Db 4264 ACAATGAAAGTACAGAGAGTAACTACTCAGGAGAGTAAATCTGCTTACAGCATGTAT 4323
QY 4215 GTGAAGAGAGTCTCATCTCCAGCAGTCAAGATGGCCATGAGTGAGCTCTTTTTCAG 4274
Db 4324 GTGAAGAGAGTCTCATCTCCAGCAGTCAAGATGGCCATGAGTGAGCTCTTTTTCAG 4383
QY 4275 AATGCCAAAGTAAAGGTTTTCAGGGAATCAAGACTCCTTCACACCTGTGTAACCTCT 4334
Db 4384 AATGCCAAAGTAAAGGTTTTCAGGGAATCAAGACTCCTTCACACCTGTGTAACCTCT 4443
QY 4335 CTAGACCCACCGTTACTGACTGCTTACCTTGAATTCACCCCGAGAGTGGGTGCACACAG 4394
Db 4444 CTAGACCCACCGTTACTGACTGCTTACCTTGAATTCACCCCGAGAGTGGGTGCACACAG 4503
QY 4395 ATTGCCCTGAGAGTGAAGGTTCTGGGCTGCGAGGACACAGACCTCTAC 4442
Db 4504 ATTGCCCTGAGAGTGAAGGTTCTGGGCTGCGAGGACACAGACCTCTAC 4551
```

```
RESULT 5
LOCUS 108642 4548 bp PAT 14-NOV-1994
DEFINITION Sequence 2 from Patent WO 8800831.
ACCESSION 108642
NID 9588648
VERSION 108642.1 GI:588648
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 4548)
AUTHORS Pask, M.P.
TITLE DNA SEQUENCES CODING FOR MODIFIED FACTOR VIII:C AND MODIFIED FACTOR
POLYPEPTIDES IN HIGH YIELDS
JOURNAL Patent: WO 8800831-A 2 11-FEB-1988;
FEATURES Location/Qualifiers
source 1..4548
BASE COUNT 1344 a 1004 c 1002 g 1197 t 1 others
ORIGIN
```

Query Match 83.7%; Score 4045.8; DB 5; Length 4548;
Best Local Similarity 94.8%; Pred. No. 0;

Matches 4311; Conservative 0; Mismatches 3; Indels 234; Gaps 3;

QY 129 GCCACCAGAAAGATACCTACCTGGGTGCAGTGGAACTGTCTATGGGACTATATGCAAGTAT 188
Db 1 GCCACCAGAAAGATACCTACCTGGGTGCAGTGGAACTGTCTATGGGACTATATGCAAGTAT 60
QY 189 CTCGGTGAAGCTGCGTGTGGACGCAAGATTTCTCCTAGAGTGCCTAAATCTTTCCATTC 248
Db 61 CTCGGTGAAGCTGCGTGTGGACGCAAGATTTCTCCTAGAGTGCCTAAATCTTTCCATTC 120
QY 249 AACACCTCAGTCGTGTACAAAAAGACTGTGTTGTAGAAATTCACGGATCACCTTTTCAAC 308
Db 121 AACACCTCAGTCGTGTACAAAAAGACTGTGTTGTAGAAATTCACGGATCACCTTTTCAAC 180
QY 309 ATCGCTAAGCCCAAGGCCACCTGGATGGGTCTGTAGTCTTACCATCCAGGCTGAGGTT 368
Db 181 ATCGCTAAGCCCAAGGCCACCTGGATGGGTCTGTAGTCTTACCATCCAGGCTGAGGTT 240
QY 369 TATGATACAGTGTGTCATTACACTTAAGAACATGGCTTCCATCCTGTGAGTCTTCATGCT 428
Db 241 TATGATACAGTGTGTCATTACACTTAAGAACATGGCTTCCATCCTGTGAGTCTTCATGCT 300
QY 429 GTTGGTGTATCCTACTGGAAGCTTCTGAGGGAGCTGAATATGATGATCAGACCAGTCAA 488
Db 301 GTTGGTGTATCCTACTGTGAAAGCTTCTGAGGGAGCTGAATATGATGATCAGACCAGTCAA 360
QY 489 AGGGAAGAAAGATGATAAAGTCTTCCCTGTGTGAAGCCATACATATGCTGGCAGGTC 548
Db 361 AGGGAAGAAAGATGATAAAGTCTTCCCTGTGTGAAGCCATACATATGCTGGCAGGTC 420
QY 549 CTGAAGAGAAATGTCCTCAATGGCCCTGTGACCCACTGTGCCCTTACCTACTATCTTCT 608
Db 421 CTGAAGAGAAATGTCCTCAATGGCCCTGTGACCCACTGTGCCCTTACCTACTATCTTCT 480
QY 609 CATGTGACCTGTGTAAAGACTTGAATTCAGGCTCATGTGAGGCCCTACTAGTATGATA 668
Db 481 CATGTGACCTGTGTAAAGACTTGAATTCAGGCTCATGTGAGGCCCTACTAGTATGATA 540
QY 669 GAAGGAGTCTGGCCAAAGAAAGACACAGACCTTGACAAATTTATACTATCTTTTGT 728
Db 541 GAAGGAGTCTGGCCAAAGAAAGACACAGACCTTGACAAATTTATACTTCTTTTGT 600
QY 729 GTATTGTGAAGGGAAGAAAGTGGCAGCTCAGAAACAAAGAACTCCTGTGATGAGATAG 788
Db 601 GTATTGTGAAGGGAAGAAAGTGGCAGCTCAGAAACAAAGAACTCCTGTGATGAGATAG 660
QY 789 GATGTCGATCTGCTCGGCTGCGCTAAATGACACAGATCAATGTTATGTAAACAG 848
Db 661 GATGTCGATCTGCTCGGCTGCGCTAAATGACACAGATCAATGTTATGTAAACAG 720
QY 849 TCTCTG---CCAGTCTGATTTGGATGCGACAGGAATCAGTCTATTTGGCATGTGATGA 905
Db 721 TCTCTGTCACAGTCTGATTTGGATGCGACAGGAATCAGTCTATTTGGCATGTGATGA 780
QY 906 ATGGGACCACTCCTGAAGTGCACCTCAATATTCCTCGAAGTCCACACATTTCTGTGAG 965
Db 781 ATGGGACCACTCCTGAAGTGCACCTCAATATTCCTCGAAGTCCACACATTTCTGTGAG 840
QY 966 AACCATGCGCCAGGCGTCTTGGAAATCTCGCCCAATAACTTTCTTACTGCTCAACACTC 1025
Db 841 AACCATGCGCCAGGCGTCTTGGAAATCTCGCCCAATAACTTTCTTACTGCTCAACACTC 900
QY 1026 TTGATGACCTTGACAGTTTCTACTGTTTGTGCATATCTCTCCACCAACATGATGGC 1085
Db 901 TTGATGACCTTGACAGTTTCTACTGTTTGTGCATATCTCTCCACCAACATGATGGC 960
QY 1086 ATGAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGAACCCCACTACGAATGAAAAAT 1145
Db 961 ATGAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGAACCCCACTACGAATGAAAAAT 1020
QY 1146 AATGAAGACGGGAAGACTATGATGATCTTACTGATCTGAAATGATGTGTGACAG 1205
Db 1021 AATGAAGACGGGAAGACTATGATGATCTTACTGATCTGAAATGATGTGTGACAG 1080

QY 1206 TTTGATGATGACAACCTCTCCTCTTATCCAAATTCGCTCAGTTCGCCAAGACATCCCT 1265
Db 1081 TTTGATGATGACAACCTCTCCTCTTATCCAAATTCGCTCAGTTCGCCAAGACATCCCT 1140
QY 1266 AAAACTTGGGTACATTACATTGCTGCTGAAGAGAGGACTGGGACTATGCTCCCTTAGTC 1325
Db 1141 AAAACTTGGGTACATTACATTGCTGCTGAAGAGAGGACTGGGACTATGCTCCCTTAGTC 1200
QY 1326 CTCGCCCCGATGACAGAAAGTTATAAAGTCAATATTTGAACAATGGCCCTCAGCGGATT 1385
Db 1201 CTCGCCCCGATGACAGAAAGTTATAAAGTCAATATTTGAACAATNGCCCTCAGCGGATT 1260
QY 1386 GGTAGGAAGTACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGT 1445
Db 1261 GGTAGGAAGTACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGT 1320
QY 1446 GAAGCTATTCAGCATGATCAGGAATCTGGGACCTTTACTTTATGGGGAAGTTGGAGAC 1505
Db 1321 GAAGCTATTCAGCATGATCAGGAATCTGGGACCTTTACTTTATGGGGAAGTTGGAGAC 1380
QY 1506 ACACGTGTTGATTTATTTAAGATCAAGCAAGCAGACCATATTAACATCTACCCTCACGGA 1565
Db 1381 ACACGTGTTGATTTATTTAAGATCAAGCAAGCAGACCATATTAACATCTACCCTCACGGA 1440
QY 1566 ATCAGTATGTCGCGCTTTGTATTCAGGAGATTAACCAAAAGGTGTAACATTTGAAG 1625
Db 1441 ATCAGTATGTCGCGCTTTGTATTCAGGAGATTAACCAAAAGGTGTAACATTTGAAG 1500
QY 1626 GATTTCCCAATTTCTGCCAGGAGAAATATTCAAATATTAATGGACAGTGTAGAGAT 1685
Db 1501 GATTTCCCAATTTCTGCCAGGAGAAATATTCAAATATTAATGGACAGTGTAGAGAT 1560
QY 1686 GGGCCAACTAAATCAGATCTCTCGGTGCTGACCCGCTATTTACTCTAGTTTGGTAAATG 1745
Db 1561 GGGCCAACTAAATCAGATCTCTCGGTGCTGACCCGCTATTTACTCTAGTTTGGTAAATG 1620
QY 1746 GAGAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGTGTCACAAAGATCTGTA 1805
Db 1621 GAGAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGTGTCACAAAGATCTGTA 1680
QY 1806 GATCAAGAGGAAACAGATATGTTCAGACAAAGAGAAATGTCAATCTGTTTCTGTAAATT 1865
Db 1681 GATCAAGAGGAAACAGATATGTTCAGACAAAGAGAAATGTCAATCTGTTTCTGTAAATT 1740
QY 1866 GATGAGAACCGAAGCTGTGCTCCTCAGAGAAATATCAAGCTTTCTCCCAATCCAGCT 1925
Db 1741 GATGAGAACCGAAGCTGTGCTCCTCAGAGAAATATCAAGCTTTCTCCCAATCCAGCT 1800
QY 1926 GGAGTCAAGCTTGAGGATCCAGAGTTCGAAGCTCCCAATCATGACAGCATCAATGCG 1985
Db 1801 GGAGTCAAGCTTGAGGATCCAGAGTTCGAAGCTCCCAATCATGACAGCATCAATGCG 1860
QY 1986 TATGTTTTGATAGTTTGCAGTTGTGTCAGTTGTTGTCATGAGGTGCGATACGTGATATT 2045
Db 1861 TATGTTTTGATAGTTTGCAGTTGTGTCAGTTGTTGTCATGAGGTGCGATACGTGATATT 1920
QY 2046 CTAAGCATTGAGACAGACTGACTTCTTCTGTCTCTCTCTGATATACCTTCAAA 2105
Db 1921 CTAAGCATTGAGACAGACTGACTTCTTCTGTCTCTCTCTGATATACCTTCAAA 1980
QY 2106 CACAAAATGCTCTATGAAGACACACTCAACCTATTCCTCTCAGGAGAAACTGTCTTC 2165
Db 1981 CACAAAATGCTCTATGAAGACACACTCAACCTATTCCTCTCAGGAGAAACTGTCTTC 2040
QY 2166 ATGTGATGAAAAACCGAGTCTATGATTTCTGGGGTGCACAACTGAGACTTTGCGAAC 2225
Db 2041 ATGTGATGAAAAACCGAGTCTATGATTTCTGGGGTGCACAACTGAGACTTTGCGAAC 2100
QY 2226 AGAGCATGACCGCTTACTGAAGGTTTCTAGTTGTGACAAAGAACTGGTATATTAC 2285
Db 2101 AGAGCATGACCGCTTACTGAAGGTTTCTAGTTGTGACAAAGAACTGGTATATTAC 2160

QY 2286 GAGGACAGTTATGAAGATATTTTCAGCATACTGCTGAGTAAACAAATGCCATTGAAACA 2345
 |||||
 Db 2161 GAGGACAGTTATGAAGATATTTTCAGCATACTGCTGAGTAAACAAATGCCATTGAAACA 2220
 QY 2346 AGAAGCTTCTCCAG----- 2361
 |||||
 Db 2221 AGAAGCTTCTCCAGGATCTTGTGCTTGGGATTAACCACTATGTTACTCAGATACCAAAA 2280
 QY 2361 ----- 2361
 Db 2281 GAAGAGTGAATCCCAAGAGAAGTCACCAGAAAAACAGCTTTTAAGAAAAAGATACC 2340
 QY 2361 ----- 2361
 Db 2341 ATTTTGTCCCTGAACGCTTGTGAAGCAATCATGCANTAGCAGCAATTAATGAGGACAA 2400
 QY 2361 ----- 2361
 2401 AATAAGCCGAAATAGAAGTCACTGGGCAAGCAAGTAGGACTGAAAGGCTGTCTCT 2460
 QY 2361 ----- 2361
 2361 ---AACCACCACTCTTGAAGCCCATCAACGGGAATTAACCTGACTACTCTTCACTCA 2417
 |||||
 Db 2461 CAAAACCCACCACTTGAAGCCCATCAACGGGAATTAACCTGACTACTCTTCACTCA 2520
 QY 2418 GATCAAGAGGAAATGACTATGATGATACCATATCAGTTGAATGAAGAGGAAGATTTT 2477
 |||||
 Db 2521 GATCAAGAGGAAATGACTATGATGATACCATATCAGTTGAATGAAGAGGAAGATTTT 2580
 QY 2478 GACATTATGATGAGGATGAATCAGAGCCCCGACAGCTTCAAAAGAAAAACAGACAC 2537
 |||||
 Db 2581 GACATTATGATGAGGATGAATCAGAGCCCCGACAGCTTCAAAAGAAAAACAGACAC 2640
 QY 2538 TATTTTATGCTCAGTGGAGAGGCTCTGGGATTTATGGGATGAGTAGCTCCACATGTT 2597
 |||||
 Db 2641 TATTTTATGCTCAGTGGAGAGGCTCTGGGATTTATGGGATGAGTAGCTCCACATGTT 2700
 QY 2598 CTAGAACAAGGGCTCAGAGTGGAGTGTCCCTCAGTTCAGAAAGTTGTTTTCCAGAA 2657
 |||||
 Db 2701 CTAGAACAAGGGCTCAGAGTGGAGTGTCCCTCAGTTCAGAAAGTTGTTTTCCAGAA 2760
 QY 2658 TTTACTGATGGCTCTTACTCAGCCCTTATACCGTGAAGAACTAAATGAACATTTGGGA 2717
 |||||
 Db 2761 TTTACTGATGGCTCTTACTCAGCCCTTATACCGTGAAGAACTAAATGAACATTTGGGA 2820
 QY 2718 CTCTGGGGCCATATATAGAAGCAAGATGAAGATATATCATGTGTAACCTTCAGAAAT 2777
 |||||
 Db 2821 CTCTGGGGCCATATATAGAAGCAAGATGAAGATATATCATGTGTAACCTTCAGAAAT 2880
 QY 2778 CAGGCTCTGCTCCCTATTTCTTCTATTTAGCCCTTATTTCTTATGAGGAAGATCAGAG 2837
 |||||
 Db 2881 CAGGCTCTGCTCCCTATTTCTTCTATTTAGCCCTTATTTCTTATGAGGAAGATCAGAG 2940
 QY 2838 CAAGGACGAACCTAGAAAAAATTTGTCAAGCCTAATGAACCAAACTTACTTTGG 2897
 |||||
 Db 2941 CAAGGACGAACCTAGAAAAAATTTGTCAAGCCTAATGAACCAAACTTACTTTGG 3000
 QY 2898 AAAGTGAACATCATATGGAACCCACACTAAAGATGAGTTGACTGCAAAAGCCTGGCTTAT 2957
 |||||
 Db 3001 AAAGTGAACATCATATGGAACCCACACTAAAGATGAGTTGACTGCAAAAGCCTGGCTTAT 3060
 QY 2958 TTTCTGATGTTGACCTGGAAGAAAGATGTGCACTCAGGCTGATTTGACCCCTTCTGTC 3017
 |||||
 Db 3061 TTTCTGATGTTGACCTGGAAGAAAGATGTGCACTCAGGCTGATTTGACCCCTTCTGTC 3120
 QY 3018 TGCCACACTAACACACTGAACCCCTGCTCATGGAGACAAGTGACAGTACAGGAATTTGCT 3077
 |||||
 Db 3121 TGCCACACTAACACACTGAACCCCTGCTCATGGAGACAAGTGACAGTACAGGAATTTGCT 3180
 QY 3078 CTGTTT---TCACCACTTTTATGAGAACCAAAAGCTGTTACTTCACTGAAATATGGA 3134
 |||||
 Db 3181 CTGTTTTCCTCACCACTTTTATGAGAACCAAAAGCTGTTACTTCACTGAAATATGGA 3240
 QY 3135 AGAACTGACAGGCTCCCTGCAATATCCAGATGGAAGATCCACTTTTAAGAGAAATAT 3194

Db 3241 AGAACTGCAGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAGAGAAATAT 3300
 |||||
 QY 3195 CGCTTCATGCAATCAATGCGCTACATATGATGATACACTACCTGGCTAGTAATGCTCAG 3254
 |||||
 Db 3301 CGCTTCATGCAATCAATGCGCTACATATGATGATACACTACCTGGCTAGTAATGCTCAG 3360
 QY 3255 GATCAAGGATTCGATGCTATCTGCTCAGCATGGGCAAGCAATGAAGCAATCCATTTAT 3314
 |||||
 Db 3361 GATCAAGGATTCGATGCTATCTGCTCAGCATGGGCAAGCAATGAAGCAATCCATTTAT 3420
 QY 3315 CATTCAGTGACATGCTGTTCACTGTCGAAAAAAGAGAGATATAAATGCACTGTAC 3374
 |||||
 Db 3421 CATTCAGTGACATGCTGTTCACTGTCGAAAAAAGAGAGATATAAATGCACTGTAC 3480
 QY 3375 AATCTATCCAGGCTGTTTTTGGAGACGTGGAATGTTACCATCCAAAGCTGGAATTTGG 3434
 |||||
 Db 3481 AATCTATCCAGGCTGTTTTTGGAGACGTGGAATGTTACCATCCAAAGCTGGAATTTGG 3540
 QY 3435 CGGTTGAATGCTTATTTGGCGAGCAATCTACATGCTGGATGAGCACACTTTTCTGCTG 3494
 |||||
 Db 3541 CGGTTGAATGCTTATTTGGCGAGCAATCTACATGCTGGATGAGCACACTTTTCTGCTG 3600
 QY 3495 TACAGCAATAGTGTGAGACTCCCTGGAATGCTTCTGACACATTTAGATTTTCAG 3554
 |||||
 Db 3601 TACAGCAATAGTGTGAGACTCCCTGGAATGCTTCTGACACATTTAGATTTTCAG 3660
 QY 3555 ATTACGCTTCAAGCAATATGAGACGTGGGCCCCAAAGCTGGCCAGACTTCATATTTCC 3614
 |||||
 Db 3661 ATTACGCTTCAAGCAATATGAGACGTGGGCCCCAAAGCTGGCCAGACTTCATATTTCC 3720
 QY 3615 GGATCAATCAATGCTGAGACACCAAGAGGCCCTTTCTTGGATCAAGGTGATCTGTTG 3674
 |||||
 Db 3721 GGATCAATCAATGCTGAGACACCAAGAGGCCCTTTCTTGGATCAAGGTGATCTGTTG 3780
 QY 3675 GCACCAATGATTTATCAAGGACATCAAGACCAGGCTGCCGTCAAGATTTCTCAGCCTC 3734
 |||||
 Db 3781 GCACCAATGATTTATCAAGGACATCAAGACCAGGCTGCCGTCAAGATTTCTCAGCCTC 3840
 QY 3735 TACATCTCAGTTATCATCATGATATAGTCTTGAATGGAAGAGTGGCAGACTTATCGA 3794
 |||||
 Db 3841 TACATCTCAGTTATCATCATGATATAGTCTTGAATGGAAGAGTGGCAGACTTATCGA 3900
 QY 3795 GGAATTCACACTGGAACCTTAATGCTCTTTGGCAATGTGATTCATCTGGGATAAA 3854
 |||||
 Db 3901 GGAATTCACACTGGAACCTTAATGCTCTTTGGCAATGTGATTCATCTGGGATAAA 3960
 QY 3855 CACAATATTTTAAACCTCCAAATATTTGCTCGATACATCCGTTTGCACCCACATCAT 3914
 |||||
 Db 3961 CACAATATTTTAAACCTCCAAATATTTGCTCGATACATCCGTTTGCACCCACATCAT 4020
 QY 3915 AGCATTCGACACTCTTGCATGGAAGTTGATGGGCTGTGATTTAAATAGTTGCAGCATG 3974
 |||||
 Db 4021 AGCATTCGACACTCTTGCATGGAAGTTGATGGGCTGTGATTTAAATAGTTGCAGCATG 4080
 QY 3975 CCATTGGGAATGGAAGTAAGCAATATCAGATGCAAGATTTACTGCTTCATCTACTTT 4034
 |||||
 Db 4081 CCATTGGGAATGGAAGTAAGCAATATCAGATGCAAGATTTACTGCTTCATCTACTTT 4140
 QY 4035 ACCAATATGTTTGCACCTGTGCTCTTCAAAAGCTGCACCTCCCAAGGAGAGT 4094
 |||||
 Db 4141 ACCAATATGTTTGCACCTGTGCTCTTCAAAAGCTGCACCTCCCAAGGAGAGT 4200
 QY 4095 AATGCTGAGACCTCAGTGTATATCCAAAAAGAGTGGCTGCAAGTGAGCTTCCAGAAG 4154
 |||||
 Db 4201 AATGCTGAGACCTCAGTGTATATCCAAAAAGAGTGGCTGCAAGTGAGCTTCCAGAAG 4260
 QY 4155 ACAATGAAGTCAAGAGAGTACTACTCAGGAGTAAATCTCTGCTTACCAGCATGTAT 4214
 |||||
 Db 4261 ACAATGAAGTCAAGAGAGTACTACTCAGGAGTAAATCTCTGCTTACCAGCATGTAT 4320
 QY 4215 GTGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTCTTTTCAG 4274
 |||||

Db 4321 GTGAAGAGACTTCCTCATCTCCAGACAGTCAAGATGGCCATCAGTGAETCTCTTTTTCAG 4380
QY 4275 AATGGCAAGTAAGGTTTTCAGGGAATCAAGACTCCTTCACACCTGTGGTGAACCTT 4334
Db 4381 AATGGCAAGTAAGGTTTTCAGGGAATCAAGACTCCTTCACACCTGTGGTGAACCTT 4440
QY 4335 CTAGACCCACCGTTACTACTGCTGCTACCTTGAATTCAACCCCGAGAGTTGGGTGCACCAG 4394
Db 4441 CTAGACCCACCGTTACTACTGCTGCTACCTTGAATTCAACCCCGAGAGTTGGGTGCACCAG 4500
QY 4395 ATTGCCCTGAGGATGAGGTTCTGGGCTGCGAGGACAGACCTCTTAC 4442
Db 4501 ATTGCCCTGAGGATGAGGTTCTGGGCTGCGAGGACAGACCTCTTAC 4548

RESULT 6

AR003710
LOCUS AR003710 9009 bp DNA PAT 04-DEC-1998
DEFINITION Sequence 1 from patent US 5744446.
ACCESSION AR003710
NID 93964969
VERSION AR003710.1 GI:3964969
KEYWORDS

SOURCE

ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9009)
AUTHORS Lollar, J.S. and Runge, M.S.
TITLE Hybrid human/animal factor VIII
JOURNAL Patent: US 5744446-A 1 28-APR-1998;
LOCATIONS Location/Qualifiers
SOURCE 1. 9009

BASE COUNT 2853 a 1907 c 1844 g 2405 t
ORIGIN

Query Match 51.1%; Score 2468.8; DB 5; Length 9009;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2473; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2345 AAGAAGCTTCTCCAGAACCCACCAAGTCTTGAACGCCATCAACGGGAATTAACCTGTAC 2404
Db 5106 AAGGCTGTGCTCTCAAAACCCACCAAGTCTTGAACGCCATCAACGGGAATTAACCTGTAC 5165
QY 2405 TACTCTTCAGTCAGATCAAGAGGAATTAAGTATGATGATACCATATCACTTGAATGAA 2464
Db 5166 TACTCTTCAGTCAGATCAAGAGGAATTAAGTATGATGATACCATATCACTTGAATGAA 5225
QY 2465 GAAGGAAGATTTGACATTTATGATGAGATGAAATCAAGAGCCCGCAGCTTCAAAA 2524
Db 5226 GAAGGAAGATTTGACATTTATGATGAGATGAAATCAAGAGCCCGCAGCTTCAAAA 5285
QY 2525 GAAACACGACACTATTTATGCTGACGTGAGAGGCTCTGGATTATGGATGAGTAG 2584
Db 5286 GAAACACGACACTATTTATGCTGACGTGAGAGGCTCTGGATTATGGATGAGTAG 5345
QY 2585 CTCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGT 2644
Db 5346 CTCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGT 5405
QY 2645 TGTTCAGGAATTTACTGATGGCTCTTACTCAGGCCCTTATACCGTGAGAACTAAA 2704
Db 5406 TGTTCAGGAATTTACTGATGGCTCTTACTCAGGCCCTTATACCGTGAGAACTAAA 5465
QY 2705 TGAACATTTGGGACTCTCGGGCCATATATAAGACAGAGTGAAGATATATCATGTG 2764
Db 5466 TGAACATTTGGGACTCTCGGGCCATATATAAGACAGAGTGAAGATATATCATGTG 5525
QY 2765 AACTTTCAGAAATCAGGCTCTGCTCCCTATTCCTTCTATCTAGCCCTTATCTTATGA 2824
Db 5526 AACTTTCAGAAATCAGGCTCTGCTCCCTATTCCTTCTATCTAGCCCTTATCTTATGA 5585

QY 2825 GGAAGATCAGAGGCAAGAGCAGAACTAGAAAAAATTGTCAAGCCTATAGAAACCA 2884
Db 5586 GGAAGATCAGAGGCAAGAGCAGAACTAGAAAAAATTGTCAAGCCTATAGAAACCA 5645
QY 2885 AACTTACTTTTGAAGTGCACATCATATGACCCCACTAAAGATGATTGACTGCA 2944
Db 5646 AACTTACTTTTGAAGTGCACATCATATGACCCCACTAAAGATGATTGACTGCA 5705
QY 2945 AGCCTGGCTTATTTCTGATGTGACCTGAAAAAGATGTGCACCTCAGCCCTGATGG 3004
Db 5706 AGCCTGGCTTATTTCTGATGTGACCTGAAAAAGATGTGCACCTCAGCCCTGATGG 5765
QY 3005 ACCCTTCTGTGCTGCCACACTAACACACTGAACCTCTCATGGAGACAAGTGACAGT 3064
Db 5766 ACCCTTCTGTGCTGCCACACTAACACACTGAACCTCTCATGGAGACAAGTGACAGT 5825
QY 3065 ACAGGAATTTGCTCTGTTTTTCCACCATCTTTGATGAGACCAAAAGCTGACTTCACTGA 3124
Db 5826 ACAGGAATTTGCTCTGTTTTTCCACCATCTTTGATGAGACCAAAAGCTGACTTCACTGA 5885
QY 3125 AATATGAAAAAGAACTCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAA 3184
Db 5886 AATATGAAAAAGAACTCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAA 5945
QY 3185 AGAATATATCGCTTCATGCAATCAATGGCTACATATGATACACTACCTGCTTACT 3244
Db 5946 AGAATATATCGCTTCATGCAATCAATGGCTACATATGATACACTACCTGCTTACT 6005
QY 3245 AATGCTCAGGATCAAGAGATTCGATGATATCTGCTCAGCATGGGCAATGAAGACAT 3304
Db 6006 AATGCTCAGGATCAAGAGATTCGATGATATCTGCTCAGCATGGGCAATGAAGACAT 6065
QY 3305 CCATTCATTCATTTTCAGTGGACATGTGCTCCTGACGAAAAAAGAGAGTATAAAT 3364
Db 6066 CCATTCATTCATTTTCAGTGGACATGTGCTCCTGACGAAAAAAGAGAGTATAAAT 6125
QY 3365 GGCACGTGACATCTCTATCCAGGTGTTTTTGAGACAGTGAAGATGTACCATCCAAAGC 3424
Db 6126 GGCACGTGACATCTCTATCCAGGTGTTTTTGAGACAGTGAAGATGTACCATCCAAAGC 6185
QY 3425 TGAATTTGGCGGTGGAATGCCCTTATTTGGCGAGCATCATGCTGGGATGACACACT 3484
Db 6186 TGAATTTGGCGGTGGAATGCCCTTATTTGGCGAGCATCATGCTGGGATGACACACT 6245
QY 3485 TTTCTGTGTACAGCAATAAGTGTACAGCTCCCTGGGAATGGCTTGTGACACATTTAG 3544
Db 6246 TTTCTGTGTACAGCAATAAGTGTACAGCTCCCTGGGAATGGCTTGTGACACATTTAG 6305
QY 3545 AGATTTTCAGATTACAGCTTCAGGACATATGAGACAGTGGGCCCCCAAGCTGGCCAGACT 3604
Db 6306 AGATTTTCAGATTACAGCTTCAGGACATATGAGACAGTGGGCCCCCAAGCTGGCCAGACT 6365
QY 3605 TCATTTATCCGGATCAATCAATGCCCTGAGACACCAAGAGCCCTTTTCTTGATCAAGGT 3664
Db 6366 TCATTTATCCGGATCAATCAATGCCCTGAGACACCAAGAGCCCTTTTCTTGATCAAGGT 6425
QY 3665 GGATCTGTGGCACCAATGATTTATTCAGGCGCATCAAGACCCAGGGTGCCTCAGAAAGTT 3724
Db 6426 GGATCTGTGGCACCAATGATTTATTCAGGCGCATCAAGACCCAGGGTGCCTCAGAAAGTT 6485
QY 3725 CTCCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGAGTGGA 3784
Db 6486 CTCCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGAGTGGA 6545
QY 3785 GACTTATCGAGGAATTCACCTGGAACCTTAATGGTCTTTTGGCAATGTGATTCATC 3844
Db 6546 GACTTATCGAGGAATTCACCTGGAACCTTAATGGTCTTTTGGCAATGTGATTCATC 6605
QY 3845 TGGGATTAACACAAATATTTTAACCTTCAATTAATGCTCGATACATCCGTTGACACC 3904
Db 6606 TGGGATTAACACAAATATTTTAACCTTCAATTAATGCTCGATACATCCGTTGACACC 6665
QY 3905 AACTCATATATAGCATTCGACAGCACTCTTCGATGAGTGTGAGTGAATTAATAG 3964

|||||
Db 6666 AACATATATAGCATTCGACGACCTCTCCGATGGAGTTGATGGGCTGTGATTTAAATAG 6725
QY 3965 TTGCACGATGCCATGGGAATGAGAGTAAGAATATCAGATGCACAGATTACTGCTTC 4024
Db 6726 TTGCAGCATGCCATGGGAATGAGAGTAAGAATATCAGATGCACAGATTACTGCTTC 6785
QY 4025 ATCTACTTTTACCAATATGTTTGCACCTGTCTCTCTCAAAGCTCGACTTCACCTCCA 4084
Db 6786 ATCTACTTTTACCAATATGTTTGCACCTGTCTCTCTCAAAGCTCGACTTCACCTCCA 6845
QY 4085 AGGAGAGTAATGCTTGAGAGCTCAGGTGAATATCCAAAAGAGTGGCTGCAAGTGA 4144
Db 6846 AGGAGAGTAATGCTTGAGAGCTCAGGTGAATATCCAAAAGAGTGGCTGCAAGTGA 6905
QY 4145 CTTCAGAGACAAATGAAGTCAAGAGTAAGTACTACTCAGGAGTAATAATCTCTGCTTAC 4204
Db 6906 CTTCAGAGACAAATGAAGTCAAGAGTAAGTACTACTCAGGAGTAATAATCTCTGCTTAC 6965
QY 4205 CAGCATGTATGTAAGAGAGTCCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCT 4264
Db 6966 CAGCATGTATGTAAGAGAGTCCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCT 7025
QY 4265 CTTTTTTCAGAAATGGCAAGTAAGTCTTTTCAGGGAATCAAGACTCTCTCAGACCTGT 4324
Db 7026 CTTTTTTCAGAAATGGCAAGTAAGTCTTTTCAGGGAATCAAGACTCTCTCAGACCTGT 7085
QY 4325 GGTGAACCTCTCTAGAACCCACCGTACTGACTCGCTACCTTCGATTCACCCCCAGAGTTG 4384
Db 7086 GGTGAACCTCTCTAGAACCCACCGTACTGACTCGCTACCTTCGATTCACCCCCAGAGTTG 7145
QY 4385 GGTGCACACAGATGGCCCTGAGAGTGGAGTCTGTGGCTGCGAGCACAGGACTCTACTG 4444
Db 7146 GGTGCACACAGATGGCCCTGAGAGTGGAGTCTGTGGCTGCGAGCACAGGACTCTACTG 7205
QY 4445 AGGCTGGCCACTGCACGACCTGCCACTGCCCTCTCCCTCTCAGCTCCAGGGCAG 4504
Db 7206 AGGCTGGCCACTGCACGACCTGCCACTGCCCTCTCCCTCTCAGCTCCAGGGCAG 7265
QY 4505 TGTCCCTCCCTGGCTTGCCTTACTCTTGTGCTAAATCCTAGACAGACTGCTTGAAG 4564
Db 7266 TGTCCCTCCCTGGCTTGCCTTACTCTTGTGCTAAATCCTAGACAGACTGCTTGAAG 7325
QY 4565 CCTCCTGAATTAATCATCATGCTCTGATCTTTGTTGGTGGGGGCCAGAGGGTGCAT 4624
Db 7326 CCTCCTGAATTAATCATCATGCTCTGATCTTTGTTGGTGGGGGCCAGAGGGTGCAT 7385
QY 4625 CCAATTTAATTAATCTTACTTATTTCTGACGCTGCTCCAGATTACTCTCTCTCC 4684
Db 7386 CCAATTTAATTAATCTTACTTATTTCTGACGCTGCTCCAGATTACTCTCTCTCC 7445
QY 4685 AATATTAATGCAAAAAGAGTGAAGAAACCTGCATGAAGCATCTCTCCCTGAAAA 4744
Db 7446 AATATTAATGCAAAAAGAGTGAAGAAACCTGCATGAAGCATCTCTCCCTGAAAA 7505
QY 4745 GTTAGGCTCTCAGAGTCAAGCACTTCTGTGTGTAAGAAACTATGTGATAACTTTG 4804
Db 7506 GTTAGGCTCTCAGAGTCAAGCACTTCTGTGTGTAAGAAACTATGTGATAACTTTG 7565
QY 4805 AAAAAGATATTTATGATGTT 4824
Db 7566 AAAAAGATATTTATGATGTT 7585
RESULT 7
LOCUS I31901 9009 bp DNA PAT 20-DEC-1996
DEFINITION Sequence 3 from patent US 5583209.
ACCESSION I31901
NID g1822692
VERSION I31901.1 GI:1822692
KEYWORDS
SOURCE Unknown.

ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9009)
AUTHORS Lollar,J.S. and Runge,M.S.
TITLE Hybrid human/porcine factor VIII
JOURNAL Patent: US 5583209-A 3 10-DEC-1996;
FEATURES Location/Qualifiers
source 1..9009
BASE COUNT 2853 a 1907 c 1844 g 2405 t
ORIGIN
Query Match 51.1%; Score 2468.8; DB 5; Length 9009;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2473; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 2345 AAGAAGCTTCTCCAGAACCCACAGTCTTGAAGCGCATCAACGGGAATACTCGTAC 2404
Db 5106 AAGGCTGTGCTCTCAAAAACCCACAGTCTTGAAGCGCATCAACGGGAATACTCGTAC 5165
QY 2405 TACTCTTCACTCAGATCAAGAGAAATGACTATGATGATACCATATCAGTTGAATGAA 2464
Db 5166 TACTCTTCACTCAGATCAAGAGAAATGACTATGATGATACCATATCAGTTGAATGAA 5225
QY 2465 GAAGGAAGATTTTGACATTTATGATGAGAGTGAAGAAATGAGACCCCGCAGCTTTCAAA 2524
Db 5226 GAAGGAAGATTTTGACATTTATGATGAGAGTGAAGAAATGAGACCCCGCAGCTTTCAAA 5285
QY 2525 GAAAAACAGACACTATTTTATTTGCTGCACTGAGAGGCTCTGGATTTAGGATGAGTAG 2584
Db 5286 GAAAAACAGACACTATTTTATTTGCTGCACTGAGAGGCTCTGGATTTAGGATGAGTAG 5345
QY 2585 CTCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTCCCTCAGTTCAAGAAAGT 2644
Db 5346 CTCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTCCCTCAGTTCAAGAAAGT 5405
QY 2645 TGTTTTCAGGAATTTTACTGATGCTGCTCTTACTCAGCCCTTATACCGTGGAGAACTAA 2704
Db 5406 TGTTTTCAGGAATTTTACTGATGCTGCTCTTACTCAGCCCTTATACCGTGGAGAACTAA 5465
QY 2705 TGAACATTTGGGACTCTGGGGCCATATATAGAGCAGAAATTGAAGTATATCATGTT 2764
Db 5466 TGAACATTTGGGACTCTGGGGCCATATATAGAGCAGAAATTGAAGTATATCATGTT 5525
QY 2765 AACTTTAGAAATCAGGCTCTGCTCCTATTTCTTATTTCTAGCTTATTTCTATGA 2824
Db 5526 AACTTTAGAAATCAGGCTCTGCTCCTATTTCTTATTTCTAGCTTATTTCTATGA 5585
QY 2825 GGAAGATCAGAGGCAAGAGCAGAACTAGAAAAAATTGTCAAGCCTAATGAACCAA 2884
Db 5586 GGAAGATCAGAGGCAAGAGCAGAACTAGAAAAAATTGTCAAGCCTAATGAACCAA 5645
QY 2885 AACTTACTTTTGAAGAGTGAACATCATATGGCACCCACTAAGATGAGTTGACTGCAA 2944
Db 5646 AACTTACTTTTGAAGAGTGAACATCATATGGCACCCACTAAGATGAGTTGACTGCAA 5705
QY 2945 AGCCTGGGCTTATTTCTCTGATGTTGACCTGGAAGAAAGATGTGCACTCAGGCTGATGG 3004
Db 5706 AGCCTGGGCTTATTTCTCTGATGTTGACCTGGAAGAAAGATGTGCACTCAGGCTGATGG 5765
QY 3005 ACCCTTCTGTCTGCTGCACACTAACAACCTGAACCTGCTCATGGGAGACAAAGTGACAGT 3064
Db 5766 ACCCTTCTGTCTGCTGCACACTAACAACCTGAACCTGCTCATGGGAGACAAAGTGACAGT 5825
QY 3065 ACAGGAATTTGCTGCTGTTTTCACCATCTTTGATGAGACCAAAAGCTGTACTCTACCTGA 3124
Db 5826 ACAGGAATTTGCTGCTGTTTTCACCATCTTTGATGAGACCAAAAGCTGTACTCTACCTGA 5885
QY 3125 AATATGGAAGAAAGTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAA 3184
Db 5886 AATATGGAAGAAAGTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAA 5945

QY 3185 AGAATATTCGCTTCATGCAATCAATGGCTACATAATGATACACTACCTGGCTTACT 3244
|||||
Db 5946 AGAATATTCGCTTCATGCAATCAATGGCTACATAATGATACACTACCTGGCTTACT 6005
QY 3245 AATGGCTCAGATCAAAAGATTGATGTTCTGCTCAGCATGGGAGCAATGAAACAT 3304
|||||
Db 6006 AATGGCTCAGATCAAAAGATTGATGTTCTGCTCAGCATGGGAGCAATGAAACAT 6065
QY 3305 CCATTCTATTCATTTCAGTGCACATGTGTTCTACTGTACGAAAAAAGAGAGTATAAAT 3364
|||||
Db 6066 CCATTCTATTCATTTCAGTGCACATGTGTTCTACTGTACGAAAAAAGAGAGTATAAAT 6125
QY 3365 GGCACGTACAACTCTCTATCCAGGTGTTTGGAGACAGTGAATGTTACATCCAAAGC 3424
|||||
Db 6126 GGCACGTACAACTCTCTATCCAGGTGTTTGGAGACAGTGAATGTTACATCCAAAGC 6185
QY 3425 TGAATTTGGGGGTGGAATGCTTATTGGGAGCATCTACATGCTGGATGAGCACACT 3484
|||||
Db 6186 TGAATTTGGGGGTGGAATGCTTATTGGGAGCATCTACATGCTGGATGAGCACACT 6245
QY 3485 TTTTCTGTTGACAGCAATAGTGTACAGACCTCCCTGGGAATGGCTTCTGACACATTAG 3544
|||||
Db 6246 TTTTCTGTTGACAGCAATAGTGTACAGACCTCCCTGGGAATGGCTTCTGACACATTAG 6305
QY 3545 AGATTTTCAGATTACAGCTTCAGAGCAATATGAGACAGTGGGCCCCAAAGCTGCGCAGACT 3604
|||||
Db 6306 AGATTTTCAGATTACAGCTTCAGAGCAATATGAGACAGTGGGCCCCAAAGCTGCGCAGACT 6365
QY 3605 TCATTATTCGGATCAATCAATGCTTGAGACCAAGAGCCCTTTTCTTGATCAAGT 3664
|||||
Db 6366 TCATTATTCGGATCAATCAATGCTTGAGACCAAGAGCCCTTTTCTTGATCAAGT 6425
QY 3665 GGATCTGTTGGCACCAGATGATTATTCAGCGCATCAAGACCCAGGCTGCCCTCAGAAGTT 3724
|||||
Db 6426 GGATCTGTTGGCACCAGATGATTATTCAGCGCATCAAGACCCAGGCTGCCCTCAGAAGTT 6485
QY 3725 CTCAGCCTCAGATCTCTCAGTTTATCATCATGTATAGTCTTGATGGAAGAAGTGCA 3784
|||||
Db 6486 CTCAGCCTCAGATCTCTCAGTTTATCATCATGTATAGTCTTGATGGAAGAAGTGCA 6545
QY 3785 GACTTATCGAGAAATTCACCTGGAACCTTAATGCTCTTTGGCAATGTCATTCATC 3844
|||||
Db 6546 GACTTATCGAGAAATTCACCTGGAACCTTAATGCTCTTTGGCAATGTCATTCATC 6605
QY 3845 TGGGATAAACACAATATTTTAAACCCTCCAAATTATGCTCGATACATCCGTTTGACCC 3904
|||||
Db 6606 TGGGATAAACACAATATTTTAAACCCTCCAAATTATGCTCGATACATCCGTTTGACCC 6665
QY 3905 AACTCATATATGATTCGACGACTCTTCGCAATGAGTGTATGGGCTGTGATTTAAATAG 3964
|||||
Db 6666 AACTCATATATGATTCGACGACTCTTCGCAATGAGTGTATGGGCTGTGATTTAAATAG 6725
QY 3965 TTGCAGCATGCCATTGGGAATGAGAGATAAGCAATATCAGATGCACAGATTACTGCTTC 4024
|||||
Db 6726 TTGCAGCATGCCATTGGGAATGAGAGATAAGCAATATCAGATGCACAGATTACTGCTTC 6785
QY 4025 ATCCTACTTTACCAATATGTTTGGCCACCTGTCTCTCTCAAAAAGCTCGACTTACCTCCA 4084
|||||
Db 6786 ATCCTACTTTACCAATATGTTTGGCCACCTGTCTCTCTCTCAAAAAGCTCGACTTACCTCCA 6845
QY 4085 AGGAGAGTAATGCCCTGAGACCTCAGGTGAATAATCCAAAAGAGTGGCTGCAAGTGA 4144
|||||
Db 6846 AGGAGAGTAATGCCCTGAGACCTCAGGTGAATAATCCAAAAGAGTGGCTGCAAGTGA 6905
QY 4145 CTTCAGAGAACAAATGAAGTACAGAGTAATACTACTCAGGAGTAAATCTCTGCTTAC 4204
|||||
Db 6906 CTTCAGAGAACAAATGAAGTACAGAGTAATACTACTCAGGAGTAAATCTCTGCTTAC 6965
QY 4205 CAGCATGTATGGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCT 4264
|||||
Db 6966 CAGCATGTATGGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCT 7025
QY 4265 CTTTTCAGATGGCAAGTAAGGTTTTCAGGGAATCAAGACTCCTTCACACCTGT 4324

Db 7026 CTTTTCAGAAATGGCAAGTAAGGTTTTCAGGAAATCAAGACTCCTTCACACCTGT 7085
QY 4325 GGTGAACCTCTAGACCCACCGTTACTGACTCGCTACCTTCGAATTCACCCCCAGAGTTG 4384
|||||
Db 7086 GGTGAACCTCTAGACCCACCGTTACTGACTCGCTACCTTCGAATTCACCCCCAGAGTTG 7145
QY 4385 GGTGACCAATTTGCCCTGAGAGTGAAGTCTTGGGCTGGAGGACAGACCTTACTG 4444
|||||
Db 7146 GGTGACCAATTTGCCCTGAGAGTGAAGTCTTGGGCTGGAGGACAGACCTTACTG 7205
QY 4445 AGGTTGGCACTGCAGACACTGCGCCTGACCTGCGCTCCCTCCTCAGCTCCAGGGCAG 4504
|||||
Db 7206 AGGTTGGCACTGCAGACACTGCGCCTGACCTGCGCTCCCTCCTCAGCTCCAGGGCAG 7265
QY 4505 TGTCCCTCCCTGGCTTGCCTTCTACCTTTGCTAAATCCTAGACACACTGCCCTGAAG 4564
|||||
Db 7266 TGTCCCTCCCTGGCTTGCCTTCTACCTTTGCTAAATCCTAGACACACTGCCCTGAAG 7325
QY 4565 CCTCCTGAATTAATCATCATCAGTCCCTGCAATTTCTTGGTGGGGGCCAGAGGGTGCA 4624
|||||
Db 7326 CCTCCTGAATTAATCATCATCAGTCCCTGCAATTTCTTGGTGGGGGCCAGAGGGTGCA 7385
QY 4625 CCAATTTAATTAATCTCTTACCTATTTTCTGACAGCTGCTGCCAGATTACTCTCTCC 4684
|||||
Db 7386 CCAATTTAATTAATCTCTTACCTATTTTCTGACAGCTGCTGCCAGATTACTCTCTCC 7445
QY 4685 AATATACTAGGCCAAAAAGAGTGAAGAGAAACCTGCATGAAAGCATTTCTCCGTA 4744
|||||
Db 7446 AATATACTAGGCCAAAAAGAGTGAAGAGAAACCTGCATGAAAGCATTTCTCCGTA 7505
QY 4745 GTTAGGCCCTCAGAGTCAACACTTCTCTGTGTGTAGAAAACTATGTGATGAACCTTTG 4804
|||||
Db 7506 GTTAGGCCCTCAGAGTCAACACTTCTCTGTGTGTAGAAAACTATGTGATGAACCTTTG 7565
QY 4805 AAAAAAGATTTATGATGTT 4824
|||||
Db 7566 AAAAAAGATTTATGATGTT 7585

RESULT 8
163424 LOCUS 9009 bp DNA 26-SEP-1997
DEFINITION Sequence 1 from patent US 5663060.
ACCESSION 163424
NID 92480997
VERSION 163424.1 GI:2480997
KEYWORDS /
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 9009)
AUTHORS Lollar,J.S. and Runge,M.S.
TITLE Hybrid human/animal factor VIII
JOURNAL Patent: US 5663060-A 1 02-SEP-1997;
FEATURES
source location/Qualifiers
1. 9009
BASE COUNT 2853 a 1907 c 1844 g 2405 t
ORIGIN
Query Match 51.1%; Score 2468.8; DB 5; Length 9009;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2473; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2345 AAGAAGCTTCTCCAGAACCCACAGTCTTGAAAAAGCCATCAACGGGAATTAATCTGCTAC 2404
|||||
Db 5106 AAGGCTGTGCTCTCAAAACCACAGTCTTGAAAAAGCCATCAACGGGAATTAATCTGCTAC 5165
QY 2405 TACTCTTACGTCAAGTCAAGAGGAATTGACTATGATGATACCATATACAGTTGAATGAA 2464
|||||
Db 5166 TACTCTTACGTCAAGTCAAGAGGAATTGACTATGATGATACCATATACAGTTGAATGAA 5225

QY 2465 GAAGGAAGATTTTGACATTTATGATGAGATGAAAATCAGAGCCCCCGCAGCTTTCAAAA 2524
|||||
Db 5226 GAAGGAAGATTTTGACATTTATGATGAGATGAAAATCAGAGCCCCCGCAGCTTTCAAAA 5285
QY 2525 GAAAAACAGACACTATATTTATTTGCTGCGATGAGAGGCTCTGGATTATGGGATGAGTAG 2584
|||||
Db 5286 GAAAAACAGACACTATATTTATTTGCTGCGATGAGAGGCTCTGGATTATGGGATGAGTAG 5345
QY 2585 CTCGCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAGAAAGT 2644
|||||
Db 5346 CTCGCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAGAAAGT 5405
QY 2645 TGTTCACAGGAATTTACTGATGGCTCCCTTACTCAGCCCTTATACCGTGAGAACTAAA 2704
|||||
Db 5406 TGTTCACAGGAATTTACTGATGGCTCCCTTACTCAGCCCTTATACCGTGAGAACTAAA 5465
2705 TGAACATTTGGGACTCCTGGGGCCATATATAGAGCAGAAAGTTGAAGATAATATCATGTGT 2764
|||||
5466 TGAACATTTGGGACTCCTGGGGCCATATATAGAGCAGAAAGTTGAAGATAATATCATGTGT 5525
QY 2765 AACTTTCAAGAAATCAGGCTCTCGTCCCTATTCCTTCTATTTCTAGCTTATTTCTATGA 2824
|||||
Db 5526 AACTTTCAAGAAATCAGGCTCTCGTCCCTATTCCTTCTATTTCTAGCTTATTTCTATGA 5585
QY 2825 GGAAGATCAGAGGCAAGGAGCAGAACTAGAAAAAACTTTGTCAAGCCTATGAACCAG 2884
|||||
Db 5586 GGAAGATCAGAGGCAAGGAGCAGAACTAGAAAAAACTTTGTCAAGCCTATGAACCAG 5645
QY 2885 AACTTACTTTTGGAAAGTGCACATCATATGGCACCCACTAAAGATGAGTTTGACTGCAA 2944
|||||
Db 5646 AACTTACTTTTGGAAAGTGCACATCATATGGCACCCACTAAAGATGAGTTTGACTGCAA 5705
QY 2945 AGCTGGGCTTATTTCTCTGATGTTGACCTGGAAAAAGATGTGCACTCAGGCTGATGG 3004
|||||
Db 5706 AGCTGGGCTTATTTCTCTGATGTTGACCTGGAAAAAGATGTGCACTCAGGCTGATGG 5765
QY 3005 ACCCTCTCTGCTGCGCACACTAACACACTGAACCTGCTCATGGGAGACAAAGTCAGAGT 3064
|||||
Db 5766 ACCCTCTCTGCTGCGCACACTAACACACTGAACCTGCTCATGGGAGACAAAGTCAGAGT 5825
QY 3065 ACAGGAATTTGCTGTTTTTTCACCATCTTTGATGAGAGCCAAAAGCTGGTACTTCACTGA 3124
|||||
Db 5826 ACAGGAATTTGCTGTTTTTTCACCATCTTTGATGAGAGCCAAAAGCTGGTACTTCACTGA 5885
QY 3125 AAATATGAAAAAGAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAA 3184
|||||
5886 AAATATGAAAAAGAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAA 5945
QY 3185 AGGAATTTATCGCTTCATGCAATCAATGGCTACATATGATGATCACTACCTGGCTTAGT 3244
|||||
Db 5946 AGGAATTTATCGCTTCATGCAATCAATGGCTACATATGATGATCACTACCTGGCTTAGT 6005
QY 3245 AATGGCTCAGGATCAAAAGGATGATGATGCTCAGCATGGGCGAGCAATGAAAAACAT 3304
|||||
Db 6006 AATGGCTCAGGATCAAAAGGATGATGATGCTCAGCATGGGCGAGCAATGAAAAACAT 6065
QY 3305 CCATTTCTATTTCACTGAGCATGTGTTCACTGTACGAAAAAAAGAGAGATATAAAAAT 3364
|||||
Db 6066 CCATTTCTATTTCACTGAGCATGTGTTCACTGTACGAAAAAAAGAGAGATATAAAAAT 6125
QY 3365 GGCACGTACAACTCTATCCAGGTGTTTTGAGACAGTGAAGTGTACCATCCAAAGC 3424
|||||
Db 6126 GGCACGTACAACTCTATCCAGGTGTTTTGAGACAGTGAAGTGTACCATCCAAAGC 6185
QY 3425 TGGAAATTTGGGGGTGAATGCTTATTTGGCAGCATCTACATGCTGGATGAGACACACT 3484
|||||
Db 6186 TGGAAATTTGGGGGTGAATGCTTATTTGGCAGCATCTACATGCTGGATGAGACACACT 6245
QY 3485 TTTTCTGCTGACAGCAATAAGTGTGAGACTCCCTGGGAATGGCTTCTGACACATTAG 3544
|||||
Db 6246 TTTTCTGCTGACAGCAATAAGTGTGAGACTCCCTGGGAATGGCTTCTGACACATTAG 6305

QY 3545 AGATTTTCAGATTACAGCTTCAGGACCAATATGAGACAGTGGGCCCCAAAGCTGGCCAGACT 3604
|||||
Db 6306 AGATTTTCAGATTACAGCTTCAGGACCAATATGAGACAGTGGGCCCCAAAGCTGGCCAGACT 6365
QY 3605 TCATTTATTCGGATCAATCAATATGCTTGAGGACCAAGAGGCTTTTCTTGATCAAGGT 3664
|||||
Db 6366 TCATTTATTCGGATCAATCAATATGCTTGAGGACCAAGAGGCTTTTCTTGATCAAGGT 6425
QY 3665 GGATCTGTTGGACCAATGATATTTACGGGATCAAGACCCAGGGGTGCCCTCAGAAAGT 3724
|||||
Db 6426 GGATCTGTTGGACCAATGATATTTACGGGATCAAGACCCAGGGGTGCCCTCAGAAAGT 6485
QY 3725 CTCACGCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGTGGCA 3784
|||||
Db 6486 CTCACGCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGTGGCA 6545
QY 3785 GACTTATCGAGAAATTCACCTGGAACCTTAATGCTCTTTTGGCAATGEGATTATC 3844
|||||
Db 6546 GACTTATCGAGAAATTCACCTGGAACCTTAATGCTCTTTTGGCAATGEGATTATC 6605
QY 3845 TGGGATAAACACAATATTTTAAACCCTCCAAATTATGCTCGATACATCCGTTGCACCC 3904
|||||
Db 6606 TGGGATAAACACAATATTTTAAACCCTCCAAATTATGCTCGATACATCCGTTGCACCC 6665
QY 3905 AACTCATTTATGCAATTCGGCACACTCTTGGCATGAGTGGCTGTGATTTTAAATAG 3964
|||||
Db 6666 AACTCATTTATGCAATTCGGCACACTCTTGGCATGAGTGGCTGTGATTTTAAATAG 6725
QY 3965 TTGCAGCATGCGATTGGGAATGGAGAGTAAAGCAATATCAGATGCACAGATTACTGCTTC 4024
|||||
Db 6726 TTGCAGCATGCGATTGGGAATGGAGAGTAAAGCAATATCAGATGCACAGATTACTGCTTC 6785
QY 4025 ATCTACTTTTACCAATATGTTTGGCACCTGGTCTCTCTCAAAAAGCTGCACCTTCACCTCCA 4084
|||||
Db 6786 ATCTACTTTTACCAATATGTTTGGCACCTGGTCTCTCTCTCAAAAAGCTGCACCTTCACCTCCA 6845
QY 4085 AGGAGAGATTAATGCCCTGAGAGACCTCAGGTGAATATCCAAAAAGAGTGGCTGCAAGTGA 4144
|||||
Db 6846 AGGAGAGATTAATGCCCTGAGAGACCTCAGGTGAATATCCAAAAAGAGTGGCTGCAAGTGA 6905
QY 4145 CTTCAGAGAACAAATGAAGAATCACAGAGTAATACTACTCAGGGAGTAAATCTCTGCTTAC 4204
|||||
Db 6906 CTTCAGAGAACAAATGAAGAATCACAGAGTAATACTACTCAGGGAGTAAATCTCTGCTTAC 6965
QY 4205 CAGCATGTATGTGAAGAGTTCCTCATCTCCAGAGTCAAGATGGCCATCAGTGGACTCT 4264
|||||
Db 6966 CAGCATGTATGTGAAGAGTTCCTCATCTCCAGAGTCAAGATGGCCATCAGTGGACTCT 7025
QY 4265 CTTTTCAGAAATGGCAAGTAAAGTTTTTTCAGGGAATCAAGACTCCTTCACACTGT 4324
|||||
Db 7026 CTTTTCAGAAATGGCAAGTAAAGTTTTTTCAGGGAATCAAGACTCCTTCACACTGT 7085
QY 4325 GGTGAACCTCTAGACCCACCGTTACTGACTCGCTACCTTGAATTCACCCCAAGTGT 4384
|||||
Db 7086 GGTGAACCTCTAGACCCACCGTTACTGACTCGCTACCTTGAATTCACCCCAAGTGT 7145
QY 4385 GGTGACAGATTTGCCCTGAGATGAGGTTCTGGGCTGGAGGCACAGACCTCTACTG 4444
|||||
Db 7146 GGTGACAGATTTGCCCTGAGATGAGGTTCTGGGCTGGAGGCACAGACCTCTACTG 7205
QY 4445 AGGTGGCCACTGAGACACCTGCACTGCGCTCACTCTCCCTCAGCTCCAGGGCAG 4504
|||||
Db 7206 AGGTGGCCACTGAGACACCTGCACTGCGCTCACTCTCCCTCAGCTCCAGGGCAG 7265
QY 4505 TGTCCCTCCCTGGCTTGCTTCACTTTGTGCTAAATCTTAGCAGACACTGCCCTTGAAG 4564
|||||
Db 7266 TGTCCCTCCCTGGCTTGCTTCACTTTGTGCTAAATCTTAGCAGACACTGCCCTTGAAG 7325
QY 4565 CCTCTGAATTAACATATCATCATGCTCATTTTGGTGGGGGCCAGAGAGGTGCAT 4624
|||||
Db 7326 CCTCTGAATTAACATATCATCATGCTCATTTTGGTGGGGGCCAGAGAGGTGCAT 7385
QY 4625 CCAATTTAACTTAACCTTACCTATTTTCTGCACTGCTCCAGATTACTCTCTCTCC 4684

|||||
Db 7386 CCAATTACTTAACCTCTTACCTATTCTGACGCTGCTCCAGATTACTCTTC
QY 4685 AATATACTAGGCAAAAAGAGTGAAGAGAAACCTGCATGAAGCATTCCTGAAAA 4744
Db 7446 AATATACTAGGCAAAAAGAGTGAAGAGAAACCTGCATGAAGCATTCCTGAAAA 7505
QY 4745 GTTAGGCTCTCAGAGTCAACCTCTCTGTGTAGAAAAAATATGTAGTGAACCTTTG 4804
Db 7506 GTTAGGCTCTCAGAGTCAACCTCTCTGTGTAGAAAAAATATGTAGTGAACCTTTG 7565
QY 4805 AAAAGATATTATGATGT 4824
Db 7566 AAAAGATATTATGATGT 7585

RESULT 9
LOCUS I71409 8967 bp DNA PAT 23-DEC-1997
DEFINITION Sequence 1 from patent US 5681746.
ACCESSION I71409
VERSION 93007544
KEYWORDS I71409.1 GI:3007544
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 8967)
AUTHORS Bodner,M., De Polo,N.J., Chang,S., Hsu,D.,Chi-Tang and Respress,J.G.
TITLE Retroviral delivery of full length factor VIII
JOURNAL Patent: US 5681746-A 1 28-OCT-1997;
FEATURES
Location/Qualifiers
1..8967
Source
BASE COUNT 2841 a 1898 c 1833 g 2395 t
ORIGIN

Query Match 51.1%; Score 2468.8; DB 5; Length 8967;
Local Similarity 99.7%; Pred. No. 0;
Matches 2473; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2825 GGAAGATCAGAGCAAGAGAGCAAGCACTAGAAAAAAGCTTTGTCAAGCCTAATGAACCA 2884
Db 5545 GGAAGATCAGAGCAAGAGAGCAAGCACTAGAAAAAAGCTTTGTCAAGCCTAATGAACCA 5604
QY 2885 AACTTACTTTTGAAGAGCAACATCATATGCGACCACCTAAAGATGAGTTTACTGCA 2944
Db 5605 AACTTACTTTTGAAGAGCAACATCATATGCGACCACCTAAAGATGAGTTTACTGCA 5664
QY 2945 AGCCTGGCCTTATTTCTCTGATGTGTGACCTGGAAGAGATGTGCACATGAGCCCTGATTGG 3004
Db 5665 AGCCTGGCCTTATTTCTCTGATGTGTGACCTGGAAGAGATGTGCACATGAGCCCTGATTGG 5724
QY 3005 ACCCCTCTGCTGCTGCCACACTAACAAGCACTGAACCCCTGCTCATGGAGACAAGTGACAGT 3064
Db 5725 ACCCCTCTGCTGCTGCCACACTAACAAGCACTGAACCCCTGCTCATGGAGACAAGTGACAGT 5784
QY 3065 ACAGGAATTTGCTCTGTTTTTTCACCACTTTTGATGAGACCAAAAGCTGTACTTCACTGA 3124
Db 5785 ACAGGAATTTGCTCTGTTTTTTCACCACTTTTGATGAGACCAAAAGCTGTACTTCACTGA 5844
QY 3125 AATATGAAAAGAAACTGCAAGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAA 3184
Db 5845 AATATGAAAAGAAACTGCAAGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAA 5904
QY 3185 AGAGAAATTTATGCTTCATGCAATCAATGCGCTACATAATGATACACTACCTGCTTAA 3244
Db 5905 AGAGAAATTTATGCTTCATGCAATCAATGCGCTACATAATGATACACTACCTGCTTAA 5964
QY 3245 AATGGCTCAGGATCAAAAGATTCGATGATATCTGCTCAGCATGGGCACAATGAAGAAAT 3304
Db 5965 AATGGCTCAGGATCAAAAGATTCGATGATATCTGCTCAGCATGGGCACAATGAAGAAAT 6024
QY 3305 CCATTCTATTCAATTCAGTGGACATGTGTTCACTGTACGAAAAAAGAGAGATATAAAT 3364
Db 6025 CCATTCTATTCAATTCAGTGGACATGTGTTCACTGTACGAAAAAAGAGAGATATAAAT 6084
QY 3365 GGCACGTACAAATCTCTATCCAGGTGTTTTTGAGACAGTGAAGATGTTACCATCCAAAGC 3424
Db 6085 GGCACGTACAAATCTCTATCCAGGTGTTTTTGAGACAGTGAAGATGTTACCATCCAAAGC 6144
QY 3425 TGAATTTGGCGGGTGAATGCTTATTTGGCGAGCATCTACATGCTGGGATGAGCACACT 3484
Db 6145 TGAATTTGGCGGGTGAATGCTTATTTGGCGAGCATCTACATGCTGGGATGAGCACACT 6204
QY 3485 TTTCTGCTGATACAGCAATAGTGTACAGCTCCCTGGGAATGCTTCTGACACATTAAG 3544
Db 6205 TTTCTGCTGATACAGCAATAGTGTACAGCTCCCTGGGAATGCTTCTGACACATTAAG 6264
QY 3545 AGATTTTCAGATTACAGCTTCAGCAATATGAGACAGTGGCCCCAAAGCTGGCCAGACT 3604
Db 6265 AGATTTTCAGATTACAGCTTCAGCAATATGAGACAGTGGCCCCAAAGCTGGCCAGACT 6324
QY 3605 TCATTATCCGGATCAATCAATGCTGAGCACCAGAGCCCTTTCTTGATCAAGGT 3664
Db 6325 TCATTATCCGGATCAATCAATGCTGAGCACCAGAGCCCTTTCTTGATCAAGGT 6384
QY 3665 GGATCTGTTGGCACCAATGATTTATTCAGGCGATCAAGACCCAGGCTGCCCTCAGAAAT 3724
Db 6385 GGATCTGTTGGCACCAATGATTTATTCAGGCGATCAAGACCCAGGCTGCCCTCAGAAAT 6444
QY 3725 CTCACGCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGATGGCA 3784
Db 6445 CTCACGCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGATGGCA 6504
QY 3785 GACTTATGAGGAATTCACACTGGAACCTTAATGTTCTTTGGCAATGTGATTCATC 3844
Db 6505 GACTTATGAGGAATTCACACTGGAACCTTAATGTTCTTTGGCAATGTGATTCATC 6564
QY 3845 TGGATAAACACAATATTTTAAACCTCAATATGCTCGATACATCCGTTTGACACC 3904
Db 6565 TGGATAAACACAATATTTTAAACCTCAATATGCTCGATACATCCGTTTGACACC 6624

QY	3905	AACTCATATATAGCATTGCGCAGCACCCTTCCGATGGAGTTGATGGCGCTGTGATTTAAATAG	3964
Db	6625	AACTCATATATAGCATTGCGCAGCACCCTTCCGATGGAGTTGATGGCGCTGTGATTTAAATAG	6684
QY	3965	TTGCAGCATGCCATTGGGGAATGGAGAGTAAGAACAATATCAGATGCACAGATTACTGCTTC	4024
Db	6685	TTGCAGCATGCCATTGGGGAATGGAGAGTAAGAACAATATCAGATGCACAGATTACTGCTTC	6744
QY	4025	ATCCTACTTTACCAATATGTTTGGCCACCCTGGTCTCCTTCAAAAAGCTCGACTTCACCTCCA	4084
Db	6745	ATCCTACTTTACCAATATGTTTGGCCACCCTGGTCTCCTTCAAAAAGCTCGACTTCACCTCCA	6804
QY	4085	AGGAGAGGTAATGCTCGGAGAGACCCTCAGCTGATATATCCAAAAAGAGTGGCTGCAAGTGA	4144
Db	6805	AGGAGAGGTAATGCTCGGAGAGACCCTCAGCTGATATATCCAAAAAGAGTGGCTGCAAGTGA	6864
QY	4145	CTTCACAAGACAATGAAGAAGTCACAGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG	4204
Db	6865	CTTCACAAGACAATGAAGAAGTCACAGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG	6924
QY	4205	CAGCATGTATGTGAAGAGAGTTCCCTCATCTCCAGCAGTCAGAGTGCATCAGTGGACTCT	4264
Db	6925	CAGCATGTATGTGAAGAGAGTTCCCTCATCTCCAGCAGTCAGAGTGCATCAGTGGACTCT	6984
QY	4265	CTTTTTCAGAATGGCAAGTAAGGTTTTTCAGGGAATCAAGACTCCTTCACACCTGT	4324
Db	6985	CTTTTTCAGAATGGCAAGTAAGGTTTTTCAGGGAATCAAGACTCCTTCACACCTGT	7044
QY	4325	GGTGAACCTCTAGAACCCACCGTTACTGACTGCTACCTTGAATTCAACCCCAAGAGTTG	4384
Db	7045	GGTGAACCTCTAGAACCCACCGTTACTGACTGCTACCTTGAATTCAACCCCAAGAGTTG	7104
QY	4385	GGTGCAACCAGATTGCCCCGAGGATGGAAGTTCTGGGCTGCGAGGACACAGGACCTTACTG	4444
Db	7105	GGTGCAACCAGATTGCCCCGAGGATGGAAGTTCTGGGCTGCGAGGACACAGGACCTTACTG	7164
QY	4445	AGGGTGGCCACTGCAGACACCTGCCACTGCCGCTCCTCCTCAGCTCCAGGCGAG	4504
Db	7165	AGGGTGGCCACTGCAGACACCTGCCACTGCCGCTCCTCCTCAGCTCCAGGCGAG	7224
QY	4505	TGTCCCTCCCTGGCTTGGCTTCTTACCTTTGTGCTAAATCCTAGCAGACACTGCCCTTGAAG	4564
Db	7225	TGTCCCTCCCTGGCTTGGCTTCTTACCTTTGTGCTAAATCCTAGCAGACACTGCCCTTGAAG	7284
QY	4565	CCCTCGAATTAATCATCATCAGTCCCTGCATTTCTTGGTGGGGGGCCAGAGGGTGCAT	4624
Db	7285	CCCTCGAATTAATCATCATCAGTCCCTGCATTTCTTGGTGGGGGGCCAGAGGGTGCAT	7344
QY	4625	CCAATTTAATTAATCTTACTTATTTTCTGACGCTGCTCCAGATTACTCTTCCCTTCC	4684
Db	7345	CCAATTTAATTAATCTTACTTATTTTCTGACGCTGCTCCAGATTACTCTTCCCTTCC	7404
QY	4685	AAATATACTAGGCAAAAAGAGTGAGAGAAAACCTGCATGAAGAATTTCTCCCTGAAA	4744
Db	7405	AAATATACTAGGCAAAAAGAGTGAGAGAAAACCTGCATGAAGAATTTCTCCCTGAAA	7464
QY	4745	GTTAGGCTCTCAGAGTCACCACTTCTCTGTGTAGAAAAACTATGTGATGAACCTTG	4804
Db	7465	GTTAGGCTCTCAGAGTCACCACTTCTCTGTGTAGAAAAACTATGTGATGAACCTTG	7524
QY	4805	AAAAAGATATTTATGATGT 4824	
Db	7525	AAAAAGATATTTATGATGT 7544	
RESULT 10			
HUMFVII			
LOCUS	HUMFVII	9009 bp	mRNA
DEFINITION	Human coagulation factor VIII:C (anti-hemophilic factor) mRNA.		
ACCESSION	K01740		
NID	9182802		
VERSION	K01740.1 GI:182802		
KEYWORDS	factor VIII; hemophilia.		

Db 7146 GGTGCACACAGATTGCCCTGAGGATGAGGTTCTGGGCTGCGAGGACAGACCTCTACTG 7205
QY 4445 AGGGTGGCCACTGACGACCTGGCCACTGCGCTGACCTCTCCCTCTCAGCTCCAGGGCAG 4504
Db 7206 AGGGTGGCCACTGACGACCTGGCCACTGCGCTGACCTCTCCCTCTCAGCTCCAGGGCAG 7265
QY 4505 TGTCCCTCCCTGGGCTTGGCTTCTACCTTTGTGCTAAATCCTAGCAGACACTGCTTGAAG 4564
Db 7266 TGTCCCTCCCTGGGCTTGGCTTCTACCTTTGTGCTAAATCCTAGCAGACACTGCTTGAAG 7325
QY 4565 CCTCCTGAATTAATCATCATGCTCCTGCATTTCTTTGGGGGGGGCCAGAGGGTGCAT 4624
Db 7326 CCTCCTGAATTAATCATCATGCTCCTGCATTTCTTTGGGGGGGGCCAGAGGGTGCAT 7385
Db 7386 CCAATTTAACCTTAACCTTAACCTTAATTTCTGCAAGCTGCTCCAGATTACTCTCTCCCTCC 7445
4685 AATATACTAGGCAAAAGAGAGAGAAACCTGCATGAAAGCATTCCTCCCTGAAA 4744
Db 7446 AATATACTAGGCAAAAGAGAGAGAAACCTGCATGAAAGCATTCCTCCCTGAAA 7505
QY 4745 GTTAGGCCCTCTCAGAGTACACCACTCTCTGTGTGTAGAAAACTATGTGATGAACCTTG 4804
Db 7506 GTTAGGCCCTCTCAGAGTACACCACTCTCTGTGTGTAGAAAACTATGTGATGAACCTTG 7565
QY 4805 AAAAGATATTATGATGTT 4824
Db 7566 AAAAGATATTATGATGTT 7585

RESULT 11
HUMFVIII
LOCUS HUMFVIII 9029 bp mRNA PRI 08-NOV-1994
DEFINITION Human coagulation factor VIII:C mRNA, complete cds.
ACCESSION M14113
NID g182817
VERSION M14113.1 GI:182817
KEYWORDS coagulation factor VIII:C.
SOURCE Human kidney, cDNA to mRNA, clones pF8[100,102]; and cell line GM1416 DNA, clone pF8-4.3.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 9029)
Truett,M.A., Blacher,R.W., Burke,R.L., Caput,D., Chu,C., Dina,D., Hartog,K., Kuo,C.H., Maslarz,F.R., Merryweather,J.P., Najarian,R., Pachl,C., Potter,S.J., Puma,J., Quiroga,M., Rall,L.B., Randolph,A., Urdea,M.S., Valenzuela,P., Dahl,H.-H.M., Favalaro,J., Hansen,J., Nordfang,O. and Ezban,M.
Characterization of the polypeptide composition of human factor VIII:C and the nucleotide sequence and expression of the human kidney cDNA
JOURNAL DNA 4 (5), 333-349 (1985)
MEDLINE 86081164
REFERENCE 2 (sites)
AUTHORS Youssoufian,H., Wong,C., Aronis,S., Platokoukis,H., Kazazian,H.H., Jr. and Antonarakis,S.E.
TITLE Moderately severe hemophilia A resulting from Glu---Gly substitution in exon 7 of the factor VIII gene
JOURNAL Am. J. Hum. Genet. 42 (6), 867-871 (1988)
MEDLINE 88220354
REFERENCE 3 (sites)
AUTHORS Bernardi,F., Legnani,C., Volinia,S., Patracchini,P., Rodorigo,G., DeRosa,V. and Marchetti,G.
TITLE A HindIII RFLP and a gene lesion in the coagulation factor VIII gene
JOURNAL Hum. Genet. 78 (4), 359-362 (1988)
MEDLINE 88197150
COMMENT [3] sites; mutation causing hemophilia.
[2] sites; mutations causing hemophilia.
Draft entry and clean copy sequence for [1] kindly provided by M.A.Truett, 26-FEB-1986.

FEATURES
source
The mutation at position 1042 results in a change of Glu to Lys, and the one at position 1043 in a change from Glu to Gly. The mutation at position 1042 produces a premature stop codon.
location/Qualifiers
1. 9029
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="Xq28"
1. 9029
/note="F-VIII mRNA"
172. 7227
/gene="F8C"
172. 228
/gene="F8C"
/note="coagulation factor VIII:C signal peptide"
172. 7227
/gene="F8C"
/note="preprocoagulation factor VIII:C"
/codon_start=1
/db_xref="GDB:G00-119-124"
/protein_id="AAA52485.1"
/db_xref="PIR:g182818"
/db_xref="GI:182818"

mat_peptide
mutation
mutation
mutation
/gene="F8C"
/note="g in wt; a in hemophiliac [2]"
1043
/gene="F8C"
/note="a in wt; g in hemophiliac [2]"

mutation	6853	/gene="F8C"	/note="c in wt; t in hemophillic [3]"
BASE COUNT	2860 a	1910 c	1848 g 2411 t
ORIGIN	185 bp upstream of SacI site.		

Query Match	51.1%;	Score 2468.8;	DB 9;	Length 9029;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 2473; Conservative	0;	Mismatches	7;	Indels 0; Gaps 0;

QY	2345	AAGAAGCTTCTCCACGAACCCACCAGTCTTTGAAACGCCATCAACGGGAAATTA	CTCGTAC	2404
Db	5127	AAGCGTGTGCTCTCAAAAACCCACCAGTCTTTGAAACGCCATCAACGGGAAATTA	CTCGTAC	5186
QY	2405	TACTCTTCAGTCAGATCAAGAGAAATTGACTATGATGATACCATATCAGTTGA	ATGAA	2464
Db	5187	TACTCTTCAGTCAGATCAAGAGAAATTGACTATGATGATACCATATCAGTTGA	ATGAA	5246
QY	2465	GAAGGAAGATTGTGACATTTATGATGAGGATGAAATCAGAGCCCCCGACGCTT	TTCAAA	2524
QY	5247	GAAGGAAGATTGTGACATTTATGATGAGGATGAAATCAGAGCCCCCGACGCTT	TTCAAA	5306
Db	5307	GAAGGAAGATTGTGACATTTATGCTGCAGTGGAGAGGCTCTGGGATTATGGG	ATGAGTAG	5366
QY	2585	CTCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCA	AGAAAGT	2644
Db	5367	CTCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCA	AGAAAGT	5426
QY	2645	TGTTTTCCAGGAATTTACTGATGGCTCTTACTCAGCCCTTATACCGTGGAGAA	CTAAA	2704
Db	5427	TGTTTTCCAGGAATTTACTGATGGCTCTTACTCAGCCCTTATACCGTGGAGAA	CTAAA	5486
QY	2705	TGACATTTTGGGACTCTGGGGCATATATAAGACGAGAAGTTGAAGATAATCAT	GAT	2764
Db	5487	TGACATTTTGGGACTCTGGGGCATATATAAGACGAGAAGTTGAAGATAATCAT	GAT	5546
QY	2765	AACTTTCAGAAATCAGGCCCTCTGCTCCCTATTCCTTCTATTCAGCCCTATTT	CTTATGA	2824
Db	5547	AACTTTCAGAAATCAGGCCCTCTGCTCCCTATTCCTTCTATTCAGCCCTATTT	CTTATGA	5606
QY	2825	GGAAGATCAGAGGCAAGGAGCAGAACCTAGAAAAAACTTTGTCAAGCCTAATGA	AAACCAA	2884
Db	5607	GGAAGATCAGAGGCAAGGAGCAGAACCTAGAAAAAACTTTGTCAAGCCTAATGA	AAACCAA	5666
QY	2885	AACTTACTTTTGGAAAGTGCACATCATATGGCACCCACTAAAGATGAGTTGACT	GCAA	2944
Db	5667	AACTTACTTTTGGAAAGTGCACATCATATGGCACCCACTAAAGATGAGTTGACT	GCAA	5726
QY	2945	AGCGTGGGCTATTTCCTCTGATGTTGACCTGGAAAAAGATGTGCACCTCAGGC	CTGATTGG	3004
Db	5727	AGCGTGGGCTATTTCCTCTGATGTTGACCTGGAAAAAGATGTGCACCTCAGGC	CTGATTGG	5786
QY	3005	ACCCCTTCTGCTGTGCCACACCTAACACACTGAACCCCTGCTCATGGGAGACA	AGTGCAGT	3064
Db	5787	ACCCCTTCTGCTGTGCCACACCTAACACACTGAACCCCTGCTCATGGGAGACA	AGTGCAGT	5846
QY	3065	ACAGGAATTTGCTCTGTTTTTTCACCATCTTTGATGAGACCAAAAGCTGGTACT	TCACCTGA	3124
Db	5847	ACAGGAATTTGCTCTGTTTTTTCACCATCTTTGATGAGACCAAAAGCTGGTACT	TCACCTGA	5906
QY	3125	AAATATGGAAGAAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACT	TTTAA	3184
Db	5907	AAATATGGAAGAAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACT	TTTAA	5966
QY	3185	AGAGAATTTATCGCTTCCATGCAATCAATGGCTACATAATGATATCACTACCTG	CTTAGT	3244
Db	5967	AGAGAATTTATCGCTTCCATGCAATCAATGGCTACATAATGATATCACTACCTG	CTTAGT	6026
QY	3245	AATGGCTCAGATCAAAAGGATTCGATGCTATCTGCTCAGCATGGGCAGCAATGA	AAACAT	3304

QY 2645 TGTTCACGGAATTTACTGATGGCTCTTACTCAGCCCTTATACCGTGAGAACTAAA 2704
|||||
Db 5365 TGTTCACGGAATTTACTGATGGCTCTTACTCAGCCCTTATACCGTGAGAACTAAA 5424
QY 2705 TGAACATTTGGACTCTGGGGCCATATATAGAGCAGAAGTTGAAGATATATCATGGT 2764
|||||
Db 5425 TGAACATTTGGACTCTGGGGCCATATATAGAGCAGAAGTTGAAGATATATCATGGT 5484
QY 2765 AACCTTCAGAAATCAGGCCCTCTGCTCCCTATTCCTTATCTAGCCTTATTTCTATGA 2824
|||||
Db 5485 AACCTTCAGAAATCAGGCCCTCTGCTCCCTATTCCTTATCTAGCCTTATTTCTATGA 5544
QY 2825 GGAAGATCAGAGCGCAGAGCAGAACTAGAAAAAATTGTCAAGCCTAATGAACCAA 2884
|||||
Db 5545 GGAAGATCAGAGCGCAGAGCAGAACTAGAAAAAATTGTCAAGCCTAATGAACCAA 5604
QY 2885 AACCTACTTTTGGAAAGTGCACATCATATGGCACCCACTAAAGATGAGTTGATGCA 2944
|||||
Db 5605 AACCTACTTTTGGAAAGTGCACATCATATGGCACCCACTAAAGATGAGTTGATGCA 5664
QY 2945 AGCTGGGCTTATTTCTCTGATGTGACCTGAAAAAGATGTGACCTCAGGCTGATGG 3004
|||||
Db 5665 AGCTGGGCTTATTTCTCTGATGTGACCTGAAAAAGATGTGACCTCAGGCTGATGG 5724
QY 3005 ACCCTCTGCTGCTCCACACTAACACACTGAACCTGCTCATGGAGACAAGTGACAGT 3064
|||||
Db 5725 ACCCTCTGCTGCTCCACACTAACACACTGAACCTGCTCATGGAGACAAGTGACAGT 5784
QY 3065 ACAGGAATTTGCTGCTGTTTTTTCACCACTTTTGAATGAGACCAAAAGCTGTACTTCACTGA 3124
|||||
Db 5785 ACAGGAATTTGCTGCTGTTTTTTCACCACTTTTGAATGAGACCAAAAGCTGTACTTCACTGA 5844
QY 3125 AAATATGGAAGAAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAA 3184
|||||
Db 5845 AAATATGGAAGAAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAA 5904
QY 3185 AGAGAATTTGCTGCTCCATGCATCAATGAGTGTACATATGATACACTACCTGGCTAGT 3244
|||||
Db 5905 AGAGAATTTGCTGCTCCATGCATCAATGAGTGTACATATGATACACTACCTGGCTAGT 5964
QY 3245 AATGCTCAGATCAAGGATTCGATGCTATCTGCTCAGCATGGCAGCAATGAACAT 3304
|||||
Db 5965 AATGCTCAGATCAAGGATTCGATGCTATCTGCTCAGCATGGCAGCAATGAACAT 6024
QY 3305 CCATCTATTCATTTCACTGAGTGTGACATGTCTCACTGTACGAAAAAGAGAGTATAAAT 3364
|||||
Db 6025 CCATCTATTCATTTCACTGAGTGTGACATGTCTCACTGTACGAAAAAGAGAGTATAAAT 6084
QY 3365 GGCACGTACAACTCTATCCAGGTGTTTTGAGACAGTGAATGTTTACCATCCAAGC 3424
|||||
Db 6085 GGCACGTACAACTCTATCCAGGTGTTTTGAGACAGTGAATGTTTACCATCCAAGC 6144
QY 3425 TGAATTTGGCGGTGGAATGCCCTTATTTGGCGAGCATCTACATGCTGGGATGAGCAGCACT 3484
|||||
Db 6145 TGAATTTGGCGGTGGAATGCCCTTATTTGGCGAGCATCTACATGCTGGGATGAGCAGCACT 6204
QY 3485 TTTCTGCTGACAGCAATAGTGTGAGTCTCCCTGGGAATGGCTTCTGACACATTAAG 3544
|||||
Db 6205 TTTCTGCTGACAGCAATAGTGTGAGTCTCCCTGGGAATGGCTTCTGACACATTAAG 6264
QY 3545 AGATTTTACAGATTACAGCTTCAGGACAATATGAGCAGTGGGCCCCCAAAAGCTGGCCAGACT 3604
|||||
Db 6265 AGATTTTACAGATTACAGCTTCAGGACAATATGAGCAGTGGGCCCCCAAAAGCTGGCCAGACT 6324
QY 3605 TCATATTTCCGGATCAATCAATGCCCTGAGACCAACCAAGGCCCTTTCTTGATCAAGGT 3664
|||||
Db 6325 TCATATTTCCGGATCAATCAATGCCCTGAGACCAACCAAGGCCCTTTCTTGATCAAGGT 6384
QY 3665 GGATCTGTGGCACCAGATGATTTATCAAGGCGATCAAGACCCAGGGTCCCTCAGAAAGTT 3724
|||||
Db 6385 GGATCTGTGGCACCAGATGATTTATTCAGGCGATCAAGACCCAGGGTCCCTCAGAAAGTT 6444
QY 3725 CTCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGAGTGCA 3784

Db 6445 CTCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGTGCA 6504
QY 3785 GACTTATCGAGAAATTCACCTGGAACCTTAATGCTCTTCTTTGGCAATGTGATTCATC 3844
|||||
Db 6505 GACTTATCGAGAAATTCACCTGGAACCTTAATGCTCTTCTTTGGCAATGTGATTCATC 6564
QY 3845 TGGATTAACACAAATATTTTAACCCCTCCAATTAATGCTCGATACATCCGTTTGCACC 3904
|||||
Db 6565 TGGATTAACACAAATATTTTAACCCCTCCAATTAATGCTCGATACATCCGTTTGCACC 6624
QY 3905 AACTCATTTAGCATTCGACGACTCTTCGATGGAAGTGTATGGGCTGTGATTTAAATAG 3964
|||||
Db 6625 AACTCATTTAGCATTCGACGACTCTTCGATGGAAGTGTATGGGCTGTGATTTAAATAG 6684
QY 3965 TTGACGATGCCATTTGGGAATGGAGAGTAAAGCAATATCAGATGCACAGATTACTGCTTC 4024
|||||
Db 6685 TTGACGATGCCATTTGGGAATGGAGAGTAAAGCAATATCAGATGCACAGATTACTGCTTC 6744
QY 4025 ATCTACTTTACCAATATGTTTGGCACCTGCTCTCTCAAAAGCTCGACTTCACTCCA 4084
|||||
Db 6745 ATCTACTTTACCAATATGTTTGGCACCTGCTCTCTCAAAAGCTCGACTTCACTCCA 6804
QY 4085 AGGAGGAGTAATGCCCTGGAGACCTCAGGTGAATATCCAAAAAGTGGCTGCAAGTGA 4144
|||||
Db 6805 AGGAGGAGTAATGCCCTGGAGACCTCAGGTGAATATCCAAAAAGTGGCTGCAAGTGA 6864
QY 4145 CTTCAGAGACAAATGAAAGTACAGAGGAGTAATCTCAGGGAGTAAATCTCTGCTTAC 4204
|||||
Db 6865 CTTCAGAGACAAATGAAAGTACAGAGGAGTAATCTCAGGGAGTAAATCTCTGCTTAC 6924
QY 4205 CAGCATGTATGTGAAGGAGTTCTCTATCTCCAGCAGTCAAGATGGCCATCAGTGCATCT 4264
|||||
Db 6925 CAGCATGTATGTGAAGGAGTTCTCTATCTCCAGCAGTCAAGATGGCCATCAGTGCATCT 6984
QY 4265 CTTTTTTCAGAAATGGCAAGTAAAGGTTTTTTCAGGGAAATCAAGACTCCTTACACCTGT 4324
|||||
Db 6985 CTTTTTTCAGAAATGGCAAGTAAAGGTTTTTTCAGGGAAATCAAGACTCCTTACACCTGT 7044
QY 4325 GGTAACCTCTTAGACCCACCGTTACTGACTGCTACCTTGAATTCACCCCCAGAGTTG 4384
|||||
Db 7045 GGTAACCTCTTAGACCCACCGTTACTGACTGCTACCTTGAATTCACCCCCAGAGTTG 7104
QY 4385 GGTGACCAAGATGGCCCTGAGATGAGGTTCTGGGCTGCGAGCAGAGGACCTCTACTG 4444
|||||
Db 7105 GGTGACCAAGATGGCCCTGAGATGAGGTTCTGGGCTGCGAGCAGAGGACCTCTACTG 7164
QY 4445 AGGCTGACCACTGAGCACCCTGCACTGCGCTCACTCTCCCTCAGCTCCAGGGCAG 4504
|||||
Db 7165 AGGCTGACCACTGAGCACCCTGCACTGCGCTCACTCTCCCTCAGCTCCAGGGCAG 7224
QY 4505 TGTCCCTCCCTGGCTTGCTTCACTTTGTGCTAAATCCTAGCAGACACTGCTTGAAG 4564
|||||
Db 7225 TGTCCCTCCCTGGCTTGCTTCACTTTGTGCTAAATCCTAGCAGACACTGCTTGAAG 7284
QY 4565 CTTCTGAATTAATCATATCATCAGTCTGCACTTTCTTGGTGGGGGCCAGAGGGTGCAT 4624
|||||
Db 7285 CTTCTGAATTAATCATATCATCAGTCTGCACTTTCTTGGTGGGGGCCAGAGGGTGCAT 7344
QY 4625 CCAATTTAATTAATCTTACCTAATTTTCTGACAGTGTCCAGATTAATCTCTCTTCC 4684
|||||
Db 7345 CCAATTTAATTAATCTTACCTAATTTTCTGACAGTGTCCAGATTAATCTCTCTTCC 7404
QY 4685 AATATTAATAGCAAAAAAGAGTGAAGAAACCTGCATGAAGCAATCTTCCCTGA AAA 4744
|||||
Db 7405 AATATTAATAGCAAAAAAGAGTGAAGAAACCTGCATGAAGCAATCTTCCCTGA AAA 7464
QY 4745 GTTAGGCTCTCAGAGTCACACATCTCTGTGTAGAAAAAACTATGTGATGAACCTTG 4804
|||||
Db 7465 GTTAGGCTCTCAGAGTCACACATCTCTGTGTAGAAAAAACTATGTGATGAACCTTG 7524
QY 4805 AAAAAATATTTATGATGTT 4824
|||||

Db 7525 AAAAGATATTATGATGTT 7544

RESULT 13

A05328 8241 bp mRNA PAT 03-MAY-1993

LOCUS A05328 8241 bp mRNA PAT 03-MAY-1993

DEFINITION Synthetic Factor VIII cDNA insert of pCIB89.

ACCESSION A05328

NID 9345061

KEYWORDS A05328.1 GI:345061

SOURCE synthetic construct.

ORGANISM synthetic construct.

REFERENCE 1 (bases 1 to 8241)

JOURNAL Patent: WO 8809813-A 7 15-DEC-1988;

FEATURES location/Qualifiers

1..8241

/organism="synthetic construct"

/db_xref="taxon:32630"

7..7062

/codon_start=1

/product="Factor VIII"

/protein_id="CAA00448.1"

/db_xref="PID:9345062"

/db_xref="GI:345062"

/translation="MQELSTCFLLRFLCFESATRRYLGAVELSWDMQSLGELP

VDARFPYRPSFPEPNTSVYKTLFEVETDHLFNIAKPRPVMGLIPTIOAEVYDT

VITLKNASHPVSILHAYGVSYWKASEGEYDDOTSOEKEDDKVFEQSHTYWQYL

KENGPMSADPLCLTYSTLHVLDVKDLSGLIGALLVCREGLAKETQTLKFTILF

AVDEGKSWHSETKNSLMQDRDAASARAWPKMHTVGVNRSIPLIGCHRSVYWHV

IGMTPEVHSIFLEGHTEFLVRNHRQASLEISPIFLTAOTLMDLQGLFCHISH

QHDGMEAVKVDSCPEEPQLRMKNNEEDYDDDLTSEMMDVRFDDNSPSFIQIRS

VAKKHPKTVWHYIAAEEDMDYAPVLAAPDDRSYSQYLNGPQRIGRKYYKVRMAY

TDEFTKREAIQHESGILGPLLYGEVDTLLIFKNQASRPYNIYPHGITDVRPIYSR

RLPKGVKHLKDFPLPGELEFKYKWTYVEDPTKSDPCLTRYSSSEVNMERDLASGL

IGPLLYCKESVDQNGNMSDKRNVLFSVDENRSWLTENIOREFPNPAGVQLED

PEFOASNMHSINGYVDSLSQVCLHEVAWYILSICAQTDPLSVFSGYTFKHKMY

YEDTLLEPFSGETVEMSMENPGLWILGCHNSDFRNGMTALLKVSQDKNTGYED

SYEDISAYLLSKNNAIEBRSEFSQNSRHPSTROKQFNATIPENDIEKTDPMFAHRPM

PKIQNVSSDLMLRLQSPPTPHGLSLDLOEAKYETPSDPSGALDSNNSISEMTHF

RPQLHSGDMVTPESGQLRLNEKLGTTAATELKLDFKVSSTNNLSTIPSDNLA

AGTDNTSLGPPSMVHYDSQDITLFGKSSPLTEGGLSLSENNDSKILSEGLM

NSQESSWGNKVSSTESGLFKGKRAHPALLTKDNALFKVISISLTKNTKSNNSATNR

KTHIDGPSLLIENSPYWNQIIESDEFEKVTPLIHDRLMDKNATLRLNHSNKT

SSKNMEMVOQKKEGPIPPDAQNPDMSFFKMLFLPESARMIORTHGKNSLNSGGQPSPK

QVLSLGEPSVEGQNFLESEKNVNVGGEFTKDVGLKEMVFPSSRNFLTNLDNHN

NTHNQEKIQEIEIEKFTLLIQENVLPQIHTVTGKNPMKNLFLSTRONVEGSDGA

YAPVLQDERSLNDSTNRKHTAHFSEKGEENLEGNQTKOIVEYACTRISNT

SOQNFVQRSKRALKOFRLPLETELEKRIIVDDTSTQMSKNMKHLTPSTLQIDYNE

KEKGAITQSPSLDCITRSHSIPQANSPLPIAKVSSPISIRPIYLRVLFODNSHL

AASYRRKDSGVQESSHFLQGAKKNNLSLAITLMTGDQREVGLSATNSVTYKVV

ENTVLPKPDLPKISGRVYELLPKVHIYOKDLEPTETNSGSPGHIDYEGALOGTEGAI

KWNEANRPKVPFLRYATESSAKTIPSKLLDPLAWDNYGTQIPKEBWSQEKSPKTA

EKKKDTILSLNACESNHAIAINEGONKPEIEVTAQKQRTERLSONPPLKREHRE

ITRTLOSDQDEIDDDTISVEMKEDPDIDEDENOSPRFQKTRHYFAAVERLM

DYGMSSSPHVLNRRAQSGSVPOFKYVFOEFTDGSFTQPLRGELNEHLGIPYIRA

EVEDNIMVTFRNQASRPYSFSSLSIYEDQROGAEPKRNFKPKNETKTYFWKQVHM

APTDEDFCKAWAYFSDVDLEKDVHSLIGPLVCHTNTLNPAHROKTVQOEALFET

IFDETKSWYFTENMERNCRAPNIQMEDPTFKENYRFAHNGIYMDTLPLGLVMAQDOR

IRWYLLSGNSNENIHSIFSGHVFYRKKEEYKMAIYNLPYGFVETEMLPKAGIMR

VECLIGELHAGMSTLFLVYNSKQTPPLGMAAGHIRDFOITAGSGOYOWAPKLABHY

SGSINAMSTKEPESWIKVDLAPMITHGIKTGARKQESSLIYQFIIIMSLDKKKWO

TYRGNSTGLMVFEGVNDSSGKIKHNPNPIIARYIRLHPHTYISIRSTLRMEWGCCL

NSCSMPLGESKAISSDAOITASSYFTNMFATWSPSKARLHLOGRENAAMPQVNNKEM

LQVDFQKTMKVTVGTTQGVKSLTSMYVKEFLISSQDGHQWTLFQNGKVKVFOGNO

DSETPVNVNSIDPLILTRYLRHPOSWYHQAIALRMEVLGCEAODLY"

BASE COUNT 2625 a 1759 c 1703 g 2154 t

ORIGIN

Query Match 51.1%; Score 2467.2; DB 5; Length 8241; Best Local Similarity 99.7%; Pred. No. 0; Matches 2472; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY	2345	AAGAAGCTTCTCCAGAACCCACCAGTCTTGAAGGCCATCAACGGGAATAACTCGTAC	2404
Db	4962	AAGCGTGTCTCTCAAAACCACCAGTCTTGAAGGCCATCAACGGGAATAACTCGTAC	5021
QY	2405	TACTCTTCAGTACATCAAGAGAAATGACTATGATGATACCATATCAGTTGAAATGAA	2464
Db	5022	TACTCTTCAGTACATCAAGAGAAATGACTATGATGATACCATATCAGTTGAAATGAA	5081
QY	2465	GAGGAGATTTTGACATTTATGATGAGGATGAAATCAGAGCCCCCGAGCTTCAAAA	2524
Db	5082	GAGGAGATTTTGACATTTATGATGAGGATGAAATCAGAGCCCCCGAGCTTCAAAA	5141
QY	2525	GAAACACGACACTATTTTATGCTGACAGTGAGAGGCTCTGGATTTATGGATGAGTAG	2584
Db	5142	GAAACACGACACTATTTTATGCTGACAGTGAGAGGCTCTGGATTTATGGATGAGTAG	5201
QY	2585	CTCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGT	2644
Db	5202	CTCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGT	5261
QY	2645	TGTTTCCAGGAATTTACTGATGGCTCTTACTCAGCCCTTATACCGTGAGAACTAAA	2704
Db	5262	TGTTTCCAGGAATTTACTGATGGCTCTTACTCAGCCCTTATACCGTGAGAACTAAA	5321
QY	2705	TGAACATTTGGGACTCTGGGCCATATATAAGACAGAAAGTTGAAGATAATATCATGCT	2764
Db	5322	TGAACATTTGGGACTCTGGGCCATATATAAGACAGAAAGTTGAAGATAATATCATGCT	5381
QY	2765	AACTTTCAGAAATAGGCCCTCTGCTCCCTATTCCTTATTCAGCCTTATTTCTTATGA	2824
Db	5382	AACTTTCAGAAATAGGCCCTCTGCTCCCTATTCCTTATTCAGCCTTATTTCTTATGA	5441
QY	2825	GGAAGATCAGAGGCAAGAGCAGAACCTAGAAAAAATTGTCAAGCCTAATGAAACCAA	2884
Db	5442	GGAAGATCAGAGGCAAGAGCAGAACCTAGAAAAAATTGTCAAGCCTAATGAAACCAA	5501
QY	2885	AACTTACTTTTGAAGTGAACATCATATGACCAACCCTAAAGATGAGTTGACTGCAA	2944
Db	5502	AACTTACTTTTGAAGTGAACATCATATGACCAACCCTAAAGATGAGTTGACTGCAA	5561
QY	2945	AGCCTGGGCTTATTTCTCTGATGTTGACCTGGAAAAAGATGTGCACCTCAGGCTGATTGG	3004
Db	5562	AGCCTGGGCTTATTTCTCTGATGTTGACCTGGAAAAAGATGTGCACCTCAGGCTGATTGG	5621
QY	3005	ACCCTTCTGCTCTGCCACACTAAACACTGAAACCCTGCTCATGGAGACAAAGTGACAGT	3064
Db	5622	ACCCTTCTGCTCTGCCACACTAAACACTGAAACCCTGCTCATGGAGACAAAGTGACAGT	5681
QY	3065	ACAGGAATTTGCTCTGTTTTCACCATCTTGATGAGACCAAAAGCTGTAACCTTCACGTGA	3124
Db	5682	ACAGGAATTTGCTCTGTTTTCACCATCTTGATGAGACCAAAAGCTGTAACCTTCACGTGA	5741
QY	3125	AAATATGAAAGAAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAA	3184
Db	5742	AAATATGAAAGAAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAA	5801
QY	3185	AGAGAATTTATGCTTCCATGCAATCAATGGCTACATAATGATACACTACCTGGCTTAGT	3244
Db	5802	AGAGAATTTATGCTTCCATGCAATCAATGGCTACATAATGATACACTACCTGGCTTAGT	5861
QY	3245	AATGGCTCAGATCAAAAGATTCGATGTTCTGCTCAGATGGGACAGATGAAGAACAT	3304
Db	5862	AATGGCTCAGATCAAAAGATTCGATGTTCTGCTCAGATGGGACAGATGAAGAACAT	5921
QY	3305	CCATTCTATTCATTTACAGTGACATGTGTTCACGTGACGAAAAAAGAGAGATATAAAT	3364
Db	5922	CCATTCTATTCATTTACAGTGACATGTGTTCACGTGACGAAAAAAGAGAGATATAAAT	5981
QY	3365	GGCACTGTACAATCTCTATCCAGGTGTTTTTGAGACAGTGAAATGTTACCATCCAAAGC	3424

Db 5982 GGCACGTGATACATCTCTATCCAGGTGTTTGTAGACAGAGTGAATGTTACATCCAAAGC 6041
QY 3425 TGAATTTGGCGGGTGAATGCTTATTTGGAGCATCTACATGCTGGATGACACACT 3484
Db 6042 TGAATTTGGCGGGTGAATGCTTATTTGGAGCATCTACATGCTGGATGACACACT 6101
QY 3485 TTTTCTGTGTACAGCAATAGTGTACAGACTCCCTGGAAATGGCTTCTGACACATTAG 3544
Db 6102 TTTTCTGTGTACAGCAATAGTGTACAGACTCCCTGGAAATGGCTTCTGACACATTAG 6161
QY 3545 AGATTTTACAGATTACAGCTTCAGGACAAATATGAGACAGTGGGCCCCAAAGCTGGCCAGACT 3604
Db 6162 AGATTTTACAGATTACAGCTTCAGGACAAATATGAGACAGTGGGCCCCAAAGCTGGCCAGACT 6221
QY 3605 TCATTATTCGGATCATCATATGCTTGAGACCAAGAGCCCTTTCTTGATCAAGT 3664
Db 6222 TCATTATTCGGATCATCATATGCTTGAGACCAAGAGCCCTTTCTTGATCAAGT 6281
QY 3665 GGATCTGTGGCACCAATGATTTATTCACGGCATCAAGACCAAGGCTGCCCTCAGAAGT 3724
Db 6282 GGATCTGTGGCACCAATGATTTATTCACGGCATCAAGACCAAGGCTGCCCTCAGAAGT 6341
QY 3725 CTCACGCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGAGAAGTGCCA 3784
Db 6342 CTCACGCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGAGAAGTGCCA 6401
QY 3785 GACTTATCGAAGAAATTCACACTGGAACCTTAATGGTCTTCTTGCAATGTGATTCATC 3844
Db 6402 GACTTATCGAAGAAATTCACACTGGAACCTTAATGGTCTTCTTGCAATGTGATTCATC 6461
QY 3845 TGGGATAAAACACAATATTTTAAACCCTCCAAATTATGCTGCATACATCCGTTTGACACC 3904
Db 6462 TGGGATAAAACACAATATTTTAAACCCTCCAAATTATGCTGCATACATCCGTTTGACACC 6521
QY 3905 AACTCATTTATAGCATTCGCAGACACTCTTCGCATGGAGTGGGCTGTGATTTAAATAG 3964
Db 6522 AACTCATTTATAGCATTCGCAGACACTCTTCGCATGGAGTGGGCTGTGATTTAAATAG 6581
QY 3965 TTGCAGCATGCCATTGGGAATGAGAGTAAGCAATATCAGATGCACAGATTACTGCTTC 4024
Db 6582 TTGCAGCATGCCATTGGGAATGAGAGTAAGCAATATCAGATGCACAGATTACTGCTTC 6641
QY 4025 ATCTTACTTTACCAATATGTTTGGCCACCCTGGTCTCTTCAAAAAGCTCGACTTCACCTCCA 4084
Db 6642 ATCTTACTTTACCAATATGTTTGGCCACCCTGGTCTCTTCAAAAAGCTCGACTTCACCTCCA 6701
QY 4085 AGGAGAGATTAATGCTTGAGACCTCAGGTGAATTAATCCAAAAGAGTGGCTGCAAGTGA 4144
Db 6702 AGGAGAGATTAATGCTTGAGACCTCAGGTGAATTAATCCAAAAGAGTGGCTGCAAGTGA 6761
QY 4145 CTTCAGAGACAATGAAGTCAAGAGTAATACTACTCAGGAGTAAATCTCTGCTTAC 4204
Db 6762 CTTCAGAGACAATGAAGTCAAGAGTAATACTACTCAGGAGTAAATCTCTGCTTAC 6821
QY 4205 CAGCATGTATGTGAAGAGTCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCT 4264
Db 6822 CAGCATGTATGTGAAGAGTCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCT 6881
QY 4265 CTTTTCAGAAATGCAAAAGTAAAGTTTTCAGGGGAAATCAAGACTCTTCACACCTGT 4324
Db 6882 CTTTTCAGAAATGCAAAAGTAAAGTTTTCAGGGGAAATCAAGACTCTTCACACCTGT 6941
QY 4325 GGTGAACCTCTAGACCCACCGTTACTGACTGCTACCTTGAATTCACCCCAAGAGTTG 4384
Db 6942 GGTGAACCTCTAGACCCACCGTTACTGACTGCTACCTTGAATTCACCCCAAGAGTTG 7001
QY 4385 GGTGCACACAGATGCGCTGAGAGTGAAGTCTGCGGCTGCGAGCACAGGACCTTACTG 4444
Db 7002 GGTGCACACAGATGCGCTGAGAGTGAAGTCTGCGGCTGCGAGCACAGGACCTTACTG 7061
QY 4445 AGGGTGGCCACTGACGACCTGCGCACTGCGTCACTCTCTCCTCAGCTCAGGGGAG 4504

Db 7062 AGGGTGGCCACTGACGACACTGCGCACTGCGCTCACCTCTCCCTCAGCTCCAGGGCAG 7121
QY 4505 TGTCCCTCCCTGGCTTGCTTCTACCTTTGCTTAATCCTAGACAGACTGCTTGAAG 4564
Db 7122 TGTCCCTCCCTGGCTTGCTTCTACCTTTGCTTAATCCTAGACAGACTGCTTGAAG 7181
QY 4565 CCTCTGAATTAATCATCATGCTCCTGCAATTTCTTTGGTGGGGGCCAGAGGGTGCAT 4624
Db 7182 CCTCTGAATTAATCATCATGCTCCTGCAATTTCTTTGGTGGGGGCCAGAGGGTGCAT 7241
QY 4625 CCAATTTAATTAATCTTACCTATTTTCTGACAGCTGCTCCAGATTACTCTTCTCC 4684
Db 7242 CCAATTTAATTAATCTTACCTATTTTCTGACAGCTGCTCCAGATTACTCTTCTCC 7301
QY 4685 AATATACTAGGCAAAAAGAGTGAAGAGAAACCTGCATGAAGCATTTCTCCCTGAAA 4744
Db 7302 AATATACTAGGCAAAAAGAGTGAAGAGAAACCTGCATGAAGCATTTCTCCCTGAAA 7361
QY 4745 GTTAGGCTCTCAGAGTCAACCACTTCTGTTGTAGAAAAAATATGTGATGAACCTTTG 4804
Db 7362 GTTAGGCTCTCAGAGTCAACCACTTCTGTTGTAGAAAAAATATGTGATGAACCTTTG 7421
QY 4805 AAAAGATATTATGATGTT 4824
Db 7422 AAAAGATATTATGATGTT 7441

RESULT 14
A07042
LOCUS A07042 8241 bp mRNA PAT 24-AUG-1993
DEFINITION Artificial mRNA for factor VIII.
ACCESSION A07042
NID 9412967
VERSION A07042.1 GI:412967
KEYWORDS factor VIII.
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 8241)
AUTHORS Van Ooyen,A.J.J., Andreoli,P.M., Van Mourik,J.A. and Pannekoek,H.
TITLE Method for the preparation of proteins with factor VIII activity by
JOURNAL microbial host cells; expression vectors, host cells, antibodies
GIST-BROCADES N.V.
FEATURES
source location/Qualifiers
1..8241
/organism="synthetic construct"
/db_xref="taxon:32630"
7..7062
/codon_start=1
/transl_table=11
/product="factor VIII"
/protein_id="CAA00607.1"
/db_xref="PID:9412968"
/db_xref="GI:412968"

CDS
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

QY 4085 AGGAGAGTAATGCTTGAGACCTCAGCTGAATATCCAAAAGAGTGGCTGCAAGTGA 4144
|||||
DB 6702 AGGAGAGTAATGCTTGAGACCTCAGCTGAATATCCAAAAGAGTGGCTGCAAGTGA 6761
QY 4145 CTTCAGAGACAAATGAAGTCAACAGAGTAACTACTCAGGAGTAAATCTCTGCTTAC 4204
|||||
DB 6762 CTTCAGAGACAAATGAAGTCAACAGAGTAACTACTCAGGAGTAAATCTCTGCTTAC 6821
QY 4205 CAGCATGTATGTGAAGAGTCTCTCACTCCAGCAGTCAAGATGGCCATCAGTGAAGTCT 4264
|||||
DB 6822 CAGCATGTATGTGAAGAGTCTCTCACTCCAGCAGTCAAGATGGCCATCAGTGAAGTCT 6881
QY 4265 CTCTTTTCAGAAATGGAAGTCTTTTCAGGAAATCAAGACTCTTCCACACTGT 4324
|||||
DB 6882 CTCTTTTCAGAAATGGAAGTCTTTTCAGGAAATCAAGACTCTTCCACACTGT 6941
QY 4325 GGTGAACCTCTCTAGACCCACCGTACTGACTCGCTACCTTGAATTCACCCCCAGAGTTG 4384
|||||
DB 6942 GGTGAACCTCTCTAGACCCACCGTACTGACTCGCTACCTTGAATTCACCCCCAGAGTTG 7001
4385 GGTGAACCTCTCTAGACCCACCGTACTGAGAGTCTCTGGCTGCGAGGACAGGACCTCTACTG 4444
|||||
DB 7002 GGTGAACCTCTCTAGACCCACCGTACTGAGAGTCTCTGGCTGCGAGGACAGGACCTCTACTG 7061
QY 4445 AGGCTGGCCACTGACGACACTGCGCACTGCCCTCTCCCTCAGCTCCAGGCGAG 4504
|||||
DB 7062 AGGCTGGCCACTGACGACACTGCGCACTGCCCTCTCCCTCAGCTCCAGGCGAG 7121
QY 4505 TGTCCTCCCTGCTGCTCTCTACCTTTTGCTAAATCCTAGACAGACTGCTTGAAG 4564
|||||
DB 7122 TGTCCTCCCTGCTGCTCTCTACCTTTTGCTAAATCCTAGACAGACTGCTTGAAG 7181
QY 4565 CCTCTGAATTAATCATCATCAGTCTCTGATTTCTTGGTGGGGGCCAGGAGGCTGAT 4624
|||||
DB 7182 CCTCTGAATTAATCATCATCAGTCTCTGATTTCTTGGTGGGGGCCAGGAGGCTGAT 7241
QY 4625 CCAATTTAATTAATCATCATCAGTCTCTGATTTCTTGGTGGGGGCCAGGAGGCTGAT 4684
|||||
DB 7242 CCAATTTAATTAATCATCATCAGTCTCTGATTTCTTGGTGGGGGCCAGGAGGCTGAT 7301
QY 4685 AATATACTAGGCAAAAGAGTGAAGAGAAACCTGCATGAAGCATTTCTTCCCTGAAAA 4744
|||||
DB 7302 AATATACTAGGCAAAAGAGTGAAGAGAAACCTGCATGAAGCATTTCTTCCCTGAAAA 7361
QY 4745 GTTAGGCTCTCAGAGTCAACCACTTCTGTGTAGAAAAAATATGATGAAGCTTTG 4804
|||||
DB 7362 GTTAGGCTCTCAGAGTCAACCACTTCTGTGTAGAAAAAATATGATGAAGCTTTG 7421
QY 4805 AAAAAGATATTATGATGTT 4824
|||||
DB 7422 AAAAAGATATTATGATGTT 7441

RESULT 15
E00527 8831 bp RNA PAT 29-SEP-1997
LOCUS Complete cDNA encoding human factor VIII.
DEFINITION E00527
ACCESSION
NID 92168806
VERSION 'E00527.1 GI:2168806
KEYWORDS JP 1985243023-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 8831)
AUTHORS Danieru,J.K., Richiyaado,M.R., Goodon,A.B. and Uiriamu,A.U.
TITLE MANUFACTURE OF FUNCTIONAL VIII FACTOR
JOURNAL Patent: JP 1985243023-A 2 03-DEC-1985;
GENENTECH INC
COMMENT OS Homo sapiens
PN JP 1985243023-A/2

PD 03-DEC-1985
PF 19-APR-1985 JP 1985085295
PR 20-APR-1984 US 84 602312
PI DANIERU JIEFURRI KEIPON, RICHIIYADO MAKU ROON, PI GOODON
AREN BIHAA,
PI URIAMU AAVIN UTSUDO
PC A6IK37/04,A6IK35/12,A6IK35/74,C07H21/04,C07K13/00,C07K15/04,
PC C12N5/00,
PC C12N15/00//C12P21/00,(C12N5/00,C12R1:91);
CC strandedness: Single;
CC topology: linear;
CC hypothetical: NO;
CC anti-sense: NO;
CC *source: cell_line=CH-2;
FH *source: chromosome/segment=X-chromosome;
FH key Location/Qualifiers
FH 5'UTR 1..109
FT sig_peptide 110..166
FT mat_peptide 167..7262
FT /product='factor VIII'
FT CDS 110..7265
FT /product='factor VIII'
FT 3'UTR 7266..<8967
FT polyA_signal 8948..8953
FT polyA_site 8968.
FEATURES
source 1..8831
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 2807 a 1865 c 1800 g 2359 t
ORIGIN
Query Match 48.9%; Score 2363; DB 5; Length 8831;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 2384; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
QY 4 GAGCTAAAGATATTTTAGAGAAAGATTAACCTTTTGTCTCTCCAGTTGAACATTGTGAGC 63
|||||
DB 42 GAGCTAAAGATATTTTAGAGAAAGATTAACCTTTTGTCTCTCCAGTTGAACATTGTGAGC 101
QY 64 AATAAGTCATGCAAAATAGAGCTCTCCACCTGCTTCTTCTGTGCGCTTTTGGCATTCGCT 123
|||||
DB 102 AATAAGTCATGCAAAATAGAGCTCTCCACCTGCTTCTTCTGTGCGCTTTTGGCATTCGCT 161
QY 124 TTAGTCCACCCAGAGATATACTACCTGGGTGAGTGAAGTGTATGGAAGTATATGCAAA 183
|||||
DB 162 TTAGTCCACCCAGAGATATACTACCTGGGTGAGTGAAGTGTATGGAAGTATATGCAAA 221
QY 184 GTGATCTCGGTGAGCTGCTGTGAGCAAGATTTCTCTAGAGTGCACAAATCTTTC 243
|||||
DB 222 GTGATCTCGGTGAGCTGCTGTGAGCAAGATTTCTCTAGAGTGCACAAATCTTTC 281
QY 244 CATTCACACCTCAGTCTGTGTACAAAAGACTCTGTTGTAGAAATCAGGATCAGCTTT 303
|||||
DB 282 CATTCACACCTCAGTCTGTGTACAAAAGACTCTGTTGTAGAAATCAGGATCAGCTTT 341
QY 304 TCAACATCGCTAAGCCAAAGCCACCTGGATGGTCTGCTAGTCTTACCATCAGGCTG 363
|||||
DB 342 TCAACATCGCTAAGCCAAAGCCACCTGGATGGTCTGCTAGTCTTACCATCAGGCTG 401
QY 364 AGTTTATGATACAGTGTATTAACCTTAAGAACATGGCTTCCCATCTGTCACTTTC 423
|||||
DB 402 AGTTTATGATACAGTGTATTAACCTTAAGAACATGGCTTCCCATCTGTCACTTTC 461
QY 424 ATGCTGTGTGTATCTACTGGAAGCTTCTGAGGAGCTGAATATGATGATCAGACCA 483
|||||
DB 462 ATGCTGTGTGTATCTACTGGAAGCTTCTGAGGAGCTGAATATGATGATCAGACCA 521
QY 484 GTCAAGGAGAGAAAGATGATTAAGTCTCCCTGTGGAAGCCATACATATGTCTGGC 543
|||||

Db 522 GTCAAAGGGAGAAAGATGATATAAGTCTTCCCTGGTGGAAAGCCATACATATGTCGCGC 581
QY 544 AGGTCCTGAAGAGAGATGGTCCATGGCCCTCTGACCCACTGTGCTTACTACTCATATC 603
Db 582 AGGTCCTGAAGAGAGATGGTCCATGGCCCTCTGACCCACTGTGCTTACTACTCATATC 641
QY 604 TTTCTCATGTGGACCTGGTAAAGACTGTAATTCAGGCCCTCATGTGAGCCCTACTAGTAT 663
Db 642 TTTCTCATGTGGACCTGGTAAAGACTGTAATTCAGGCCCTCATGTGAGCCCTACTAGTAT 701
QY 664 GTAGAGAAGGAGAGTCTGGCCAAAGAAAAGACACAGACCTTGCACAAATTTATACTATT 723
Db 702 GTAGAGAAGGAGAGTCTGGCCAAAGAAAAGACACAGACCTTGCACAAATTTATACTATT 761
QY 724 TTGCTGTAATTTGATGAAGGAAAAGTTGGCAGCTCAGAAAACAAAGAACTCCTGTATGACAG 783
Db 762 TTGCTGTAATTTGATGAAGGAAAAGTTGGCAGCTCAGAAAACAAAGAACTCCTGTATGACAG 821
QY 784 ATAGGGATGCTGCATCTGCTGGGCCCTGGCCCTAAATGACACAGCTCAATGGTTATGTAA 843
Db 822 ATAGGGATGCTGCATCTGCTGGGCCCTGGCCCTAAATGACACAGCTCAATGGTTATGTAA 881
QY 844 ACAGGTCTCTGCCAGGCTCTGATGGATGCCACAGAAATCAATCTATTGGCATGTGATTG 903
Db 882 ACAGGTCTCTGCCAGGCTCTGATGGATGCCACAGAAATCAATCTATTGGCATGTGATTG 941
QY 904 GAATGGGCACCACTCTCTGAAGTGCATCAATATCTCTGAAAGTCAACATCTTCTTGTGA 963
Db 942 GAATGGGCACCACTCTCTGAAGTGCATCAATATCTCTGAAAGTCAACATCTTCTTGTGA 1001
QY 964 GGAACCATCGCCAGGCGCTCCTGGGAAATCTCGCCAATTAATCTTCTTACTGCTCAAAAC 1023
Db 1002 GGAACCATCGCCAGGCGCTCCTGGGAAATCTCGCCAATTAATCTTCTTACTGCTCAAAAC 1061
QY 1024 TCTTGATGACCTTGGACAGTTTCTACTGTTTGTCTATATCTTCCCAACATGATG 1083
Db 1062 TCTTGATGACCTTGGACAGTTTCTACTGTTTGTCTATATCTTCCCAACATGATG 1121
QY 1084 GCATGGAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGAAAGCCCACTACGAATGAAAA 1143
Db 1122 GCATGGAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGAAAGCCCACTACGAATGAAAA 1181
QY 1144 ATATGAAGAACGGGAAGACTATGATGATCTTACTGATCTGAAATGATGTGTCA 1203
Db 1182 ATATGAAGAACGGGAAGACTATGATGATCTTACTGATCTGAAATGATGTGTCA 1241
QY 1204 GGTTCGATGATGACAACTCTCTCTTATTCACAAATTCGCTCAGTGGCCAGAAAGCATC 1263
Db 1242 GGTTCGATGATGACAACTCTCTCTTATTCACAAATTCGCTCAGTGGCCAGAAAGCATC 1301
QY 1264 CTAAACTTGGGTACATTTACATTTGCTGCTGAAGAGAGAGACTGGGACTATGCTCCCTAG 1323
Db 1302 CTAAACTTGGGTACATTTGCTGCTGAAGAGAGAGACTGGGACTATGCTCCCTAG 1361
QY 1324 TCCTCGCCCCGATGACAGAAAGTTATAAAGTCAATATTTGAACAATGGCCCTCAGCGGA 1383
Db 1362 TCCTCGCCCCGATGACAGAAAGTTATAAAGTCAATATTTGAACAATGGCCCTCAGCGGA 1421
QY 1384 TTGGTAGGAAGTACAAAAAGTCCGATTTATGGCATACAGATGAAACCTTTAAGACTC 1443
Db 1422 TTGGTAGGAAGTACAAAAAGTCCGATTTATGGCATACAGATGAAACCTTTAAGACTC 1481
QY 1444 GTGAAGCTATTACAGCATGATCAGGAATCTTGGGACCTTACTTTATGGGGAAGTTGGAG 1503
Db 1482 GTGAAGCTATTACAGCATGATCAGGAATCTTGGGACCTTACTTTATGGGGAAGTTGGAG 1541
QY 1504 ACACACTGTGATTATATTTAAGAAATCAAGCAAGCAGACATATTAACATCTACCCTCAGC 1563
Db 1542 ACACACTGTGATTATATTTAAGAAATCAAGCAAGCAGACATATTAACATCTACCCTCAGC 1601
QY 1564 GAATCACTGATGTCCGCTCTTGTATTCAGAGAGATTAACCAAAAGGTGTAACATTTGA 1623
Db 1602 GAATCACTGATGTCCGCTCTTGTATTCAGAGAGATTAACCAAAAGGTGTAACATTTGA 1661

QY 1624 AGGATTTTCCAAATCTGCCAGAGAAAATATTCAAATATATAATGACAGTACTGTAGAG 1683
Db 1662 AGGATTTTCCAAATCTGCCAGAGAAAATATTCAAATATATAATGACAGTACTGTAGAG 1721
QY 1684 ATGGCCAACTAAATCAGATTCCTGGTGGCTGACCCGCTATTTACTGTAGTTTGTAAAT 1743
Db 1722 ATGGCCAACTAAATCAGATTCCTGGTGGCTGACCCGCTATTTACTGTAGTTTGTAAAT 1781
QY 1744 TGGAGAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCCCTCATCTGTACAAAGAAATCTG 1803
Db 1782 TGGAGAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCCCTCATCTGTACAAAGAAATCTG 1841
QY 1804 TAGATCAAGAGAGAAAACAGATTAATGTGACACAAGAGAAATGTATCTGTTTGTAT 1863
Db 1842 TAGATCAAGAGAGAAAACAGATTAATGTGACACAAGAGAAATGTATCTGTTTGTAT 1901
QY 1864 TTGATGAGAACCGAAGCTGGTACCTCAGAGAGAAATATACAGCGCTTCTCCCAATCCAG 1923
Db 1902 TTGATGAGAACCGAAGCTGGTACCTCAGAGAGAAATATACAGCGCTTCTCCCAATCCAG 1961
QY 1924 CTGAGTGCAGCTTGAGGATCCAGATTCAGATTCAGAGCCCTCCACATCATGACATCAATG 1983
Db 1962 CTGAGTGCAGCTTGAGGATCCAGATTCAGAGATTCAGAGCCCTCCACATCATGACATCAATG 2021
QY 1984 GCTATGTTTTGATAGTTTGCAGTTTGCAGTTTGTGTCATGAGTGGCATACTGTGACA 2043
Db 2022 GCTATGTTTTGATAGTTTGCAGTTTGCAGTTTGTGTCATGAGTGGCATACTGTGACA 2081
QY 2044 TTCTAAGCATTGAGACACAGACTGACTTCTTCTCTCTCTCTCTCTGATATACCTTCA 2103
Db 2082 TTCTAAGCATTGAGACACAGACTGACTTCTTCTCTCTCTCTCTCTGATATACCTTCA 2141
QY 2104 AACACAAAATGCTCTATGAAGACACACTCAACCTATTCOCATTTCTGAGGAGAACTGTCT 2163
Db 2142 AACACAAAATGCTCTATGAAGACACACTCAACCTATTCOCATTTCTGAGGAGAACTGTCT 2201
QY 2164 TCATGTCGATGGAAGAACCCAGGCTATAGGATTTCTGGGGTGGCCACAACCTCAGACTTTCGGA 2223
Db 2202 TCATGTCGATGGAAGAACCCAGGCTATAGGATTTCTGGGGTGGCCACAACCTCAGACTTTCGGA 2261
QY 2224 ACAGAGCATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAAAGACACTGCTGATATT 2283
Db 2262 ACAGAGCATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAAAGACACTGCTGATATT 2321
QY 2284 ACAGAGCAGTTATGAGATATTTGACATACTTCTGAGTAATAAACAATGCCATTGAAC 2343
Db 2322 ACAGAGCAGTTATGAGATATTTGACATACTTCTGAGTAATAAACAATGCCATTGAAC 2381
QY 2344 CAAGAGCTTCTCCAGAACCCACAGCTCTGAAGAGCCATCAACGGGAATTAATCTGTA 2403
Db 2382 CAAGAGCTTCTCCAGAAATTCAGAGAACACCCCTAGCACTAGGCAAAAAGCAATTTAATGCCA 2441
QY 2404 CTACTCTTCAGTCAGATCA 2422
Db 2442 CCACAAATTCAGAAAATGA 2460

Search completed: August 18, 1999, 15:12:34
Job time: 2582 sec

QY 241 TTCCATTCAACCTCAGTCGTGTACAAAAAGACTCTGTTGTAGAAATTCAGGATCACC 300
|||||
Db 241 TTCCATTCAACCTCAGTCGTGTACAAAAAGACTCTGTTGTAGAAATTCAGGATCACC 300
QY 301 TTTTCAACATCGCTAAGCCAAAGCCACCCTGGATGGGTCTGTAGGTCTTACCATCCAGG 360
|||||
Db 301 TTTTCAACATCGCTAAGCCAAAGCCACCCTGGATGGGTCTGTAGGTCTTACCATCCAGG 360
QY 361 CIGAGGTTTATGATACAGTGTGATTTACACTTAAGAAACATGGCTTCCATCCTGTCAAGT 420
|||||
Db 361 CTGAGGTTTATGATACAGTGTGATTTACACTTAAGAAACATGGCTTCCATCCTGTCAAGT 420
QY 421 TTCAATGCTGTGGTGTATCTCTACTGGAAGCTCTGAGGAGCTGAATATGATGATCAGA 480
|||||
Db 421 TTCAATGCTGTGGTGTATCTCTACTGGAAGCTCTGAGGAGCTGAATATGATGATCAGA 480
QY 481 CCACTCAAAAGGAGAAAGAAAGATGATAAAGTCTTCCCTGCTGGAAGCCATACATATGTCT 540
|||||
Db 481 CCACTCAAAAGGAGAAAGAAAGATGATAAAGTCTTCCCTGCTGGAAGCCATACATATGTCT 540
QY 541 GGCAGGTCCTGAAGAGAAATGGTCCAATGGCCTCTGACCCCACTGTGCTTACTACTCAT 600
|||||
Db 541 GGCAGGTCCTGAAGAGAAATGGTCCAATGGCCTCTGACCCCACTGTGCTTACTACTCAT 600
QY 601 ATCTTTCTCATGTGTGACCTGTGTAAGAAAGACTTGAATTCAGGCCTCATTTGAGCCCTACTAG 660
|||||
Db 601 ATCTTTCTCATGTGTGACCTGTGTAAGAAAGACTTGAATTCAGGCCTCATTTGAGCCCTACTAG 660
QY 661 TATGTAGAGAAGGAGTCTGGCCAAAGAAAAGACACAGACCTTGACAAATTTATACTAC 720
|||||
Db 661 TATGTAGAGAAGGAGTCTGGCCAAAGAAAAGACACAGACCTTGACAAATTTATACTAC 720
QY 721 TTTTGTCTGATTTTGTATGAAGGGAAAAAGTTGGCACTCAGAAAACAAGAATCCTTGATGC 780
|||||
Db 721 TTTTGTCTGATTTTGTATGAAGGGAAAAAGTTGGCACTCAGAAAACAAGAATCCTTGATGC 780
QY 781 AGGATAGGATGCTGCATCTGCTCGGCTGAGCTTAAATGCACACAGTCAATGTTATG 840
|||||
Db 781 AGGATAGGATGCTGCATCTGCTCGGCTGAGCTTAAATGCACACAGTCAATGTTATG 840
QY 841 TAAACAGGTCCTGCGCAGGTCGTGATTTGAGCCACAGAAATCACTCTATGGCATGTGA 900
|||||
Db 841 TAAACAGGTCCTGCGCAGGTCGTGATTTGAGCCACAGAAATCACTCTATGGCATGTGA 900
QY 901 TTGGAATGGGCACACACTCTCTGAAGTGCACATCAATATTCCTCGAAAGTCAACATTTCTTG 960
|||||
Db 901 TTGGAATGGGCACACACTCTCTGAAGTGCACATCAATATTCCTCGAAAGTCAACATTTCTTG 960
QY 961 TGAGGAACCATCGCCAGGCGTCTTGAAATCTCGCCAAATTAATTTCTTACTGCTCAAA 1020
|||||
Db 961 TGAGGAACCATCGCCAGGCGTCTTGAAATCTCGCCAAATTAATTTCTTACTGCTCAAA 1020
QY 1021 CACTCTTGTGAGACCTTGACAGTTTCTACTGTTTGTGATATCTCTTCCACCAACATG 1080
|||||
Db 1021 CACTCTTGTGAGACCTTGACAGTTTCTACTGTTTGTGATATCTCTTCCACCAACATG 1080
QY 1081 ATGGCATGGAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGGAACCCCACTACGAATGA 1140
|||||
Db 1081 ATGGCATGGAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGGAACCCCACTACGAATGA 1140
QY 1141 AAAATATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAATGATGTGG 1200
|||||
Db 1141 AAAATATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAATGATGTGG 1200
QY 1201 TCAGGTTGATGATGACAACTCTCCTTCTTATCCAAATTCGCTCAGTTGCCAAGAAGC 1260
|||||
Db 1201 TCAGGTTGATGATGACAACTCTCCTTCTTATCCAAATTCGCTCAGTTGCCAAGAAGC 1260
QY 1261 ATCTTAAACTTGGGTACTTACATTTGCTGCTGAAGAGAGGAGACTGGGACTATGCTCCT 1320
|||||
Db 1261 ATCTTAAACTTGGGTACTTACATTTGCTGCTGAAGAGAGGAGACTGGGACTATGCTCCT 1320

QY 1321 TAGTCTGCCCCCGATGACAGAGTATATAAAGTCAATATTTGAACAATGGCCCTCACC 1380
|||||
Db 1321 TAGTCTGCCCCCGATGACAGAGTATATAAAGTCAATATTTGAACAATGGCCCTCACC 1380
QY 1381 GGATGGTAGGAAGTACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGA 1440
|||||
Db 1381 GGATGGTAGGAAGTACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGA 1440
QY 1441 CTCGTGAAGCTATTACAGATGAATCAGAAATCTTGGACCTTACTTTATGGGAAGTTG 1500
|||||
Db 1441 CTCGTGAAGCTATTACAGATGAATCAGAAATCTTGGACCTTACTTTATGGGAAGTTG 1500
QY 1501 GAGACACACTGTTGATTTATTTAAGAAATCAAGCAAGACAGACCATATACATCTACCCTC 1560
|||||
Db 1501 GAGACACACTGTTGATTTATTTAAGAAATCAAGCAAGACAGACCATATACATCTACCCTC 1560
QY 1561 ACGGAATCACTGATGTCCGCTTGTGATTTCAAGGAGTTTACCAAAAGGTGTAAACATTT 1620
|||||
Db 1561 ACGGAATCACTGATGTCCGCTTGTGATTTCAAGGAGTTTACCAAAAGGTGTAAACATTT 1620
QY 1621 TGAAGGATTTTCCAATCTGCGCAGGAGAAATATTCAAATATAATGACAGTGCCTGTAG 1680
|||||
Db 1621 TGAAGGATTTTCCAATCTGCGCAGGAGAAATATTCAAATATAATGACAGTGCCTGTAG 1680
QY 1681 AAGATGGCCCACTAAATCAGATCTCGGTGCGCTGACCCGCTATTTACTCTAGTTGCTTA 1740
|||||
Db 1681 AAGATGGCCCACTAAATCAGATCTCGGTGCGCTGACCCGCTATTTACTCTAGTTGCTTA 1740
QY 1741 ATATGAGAGAGATCTAGCTTACAGACTCATTTGGCCCTCTCCTCATCTGCTACAAAGAA 1800
|||||
Db 1741 ATATGAGAGAGATCTAGCTTACAGACTCATTTGGCCCTCTCCTCATCTGCTACAAAGAA 1800
QY 1801 CTGTAGATCAAAAGAGAAACCAGATATGTCAGACAAAGAGAAATGTCATCCTGTTTCTG 1860
|||||
Db 1801 CTGTAGATCAAAAGAGAAACCAGATATGTCAGACAAAGAGAAATGTCATCCTGTTTCTG 1860
QY 1861 TATTTGATGAGAACCGAAGCTGTGACTCTACAGAGAAATATACAAGCCTTCTCCCAATC 1920
|||||
Db 1861 TATTTGATGAGAACCGAAGCTGTGACTCTACAGAGAAATATACAAGCCTTCTCCCAATC 1920
QY 1921 CAGCTGAGTGCAGCTTGAGATCCAGAGTTCCAAAGCCTCCAAATCATGACACAGATCA 1980
|||||
Db 1921 CAGCTGAGTGCAGCTTGAGATCCAGAGTTCCAAAGCCTCCAAATCATGACACAGATCA 1980
QY 1981 ATGGCTATGTTTTGATGATTTGAGTTGTGAGTTGTCAAGTTGTTTGCATGAGGTGGCATACTG 2040
|||||
Db 1981 ATGGCTATGTTTTGATGATTTGAGTTGTGAGTTGTCAAGTTGTTTGCATGAGGTGGCATACTG 2040
QY 2041 ACATTTCTAAGCATGTGACACAGACTGACTTCTTCTGTCTCTCTGATATACCT 2100
|||||
Db 2041 ACATTTCTAAGCATGTGACACAGACTGACTTCTTCTGTCTCTCTGATATACCT 2100
QY 2101 TCAACACAAAAATGGTCTATGAAGACACACTCACCCCTATTTCCATTTCTCAGAGAAAACTG 2160
|||||
Db 2101 TCAACACAAAAATGGTCTATGAAGACACACTCACCCCTATTTCCATTTCTCAGAGAAAACTG 2160
QY 2161 TCTTCAATGTCGATGGAATAAACCAGGTCATGATTTCTGGGTGCCACAACACTCAGACTTTC 2220
|||||
Db 2161 TCTTCAATGTCGATGGAATAAACCAGGTCATGATTTCTGGGTGCCACAACACTCAGACTTTC 2220
QY 2221 GGAACAGAGGATGACCGCCTTACTGAAGTTTCTAGTTGTGACAGAACAACACTGGTGATT 2280
|||||
Db 2221 GGAACAGAGGATGACCGCCTTACTGAAGTTTCTAGTTGTGACAGAACAACACTGGTGATT 2280
QY 2281 ATTACAGAGACAGTTATGAAGATATTTACAGATTTCTGCTGAGTAAAAACAATGGCAATTG 2340
|||||
Db 2281 ATTACAGAGACAGTTATGAAGATATTTACAGATTTCTGCTGAGTAAAAACAATGGCAATTG 2340
QY 2341 AACCAAGAAGCTTTCCCGAAGAACCCAGTCTTGAAACGCCATCAACGGGAATAACTC 2400
|||||
Db 2341 AACCAAGAAGCTTTCCCGAAGAACCCAGTCTTGAAACGCCATCAACGGGAATAACTC 2400
QY 2401 GTACTACTCTTCACTCAGATCAAGAGGAATTTGACTATGATGATACCATATACGTGAAA 2460
|||||

Db 2401 GTACTACTCTTCAGTCAGATCAAGAGAAATTGACTATGATGATACCATATCAGTTGAAA 2460
QY 2461 TGAAGAAGAGATTTTGACATTTATGATGAGATGAAATCAGAGCCCCGACGCTTTC 2520
Db 2461 TGAAGAAGAGATTTTGACATTTATGATGAGATGAAATCAGAGCCCCGACGCTTTC 2520
QY 2521 AAAAGAAAACACGACACTATTTTATGCTGCAGTGGAGAGGCTCTGGGATTATGGATGA 2580
Db 2521 AAAAGAAAACACGACACTATTTTATGCTGCAGTGGAGAGGCTCTGGGATTATGGATGA 2580
QY 2581 GTAGCTCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGA 2640
Db 2581 GTAGCTCCCCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGA 2640
QY 2641 AAGTGTTTTCCAGSAATTTACTGATGGCTCCTTACTCAGCCCTTATACCGTGGAGAAC 2700
Db 2641 AAGTGTTTTCCAGSAATTTACTGATGGCTCCTTACTCAGCCCTTATACCGTGGAGAAC 2700
QY 2701 TAAATGAACATTTGGGACTCCTGGGGCCATATATAGAGCAGAGTTGAAGATATATCA 2760
Db 2701 TAAATGAACATTTGGGACTCCTGGGGCCATATATAGAGCAGAGTTGAAGATATATCA 2760
QY 2761 TGGTAACCTTTCAGAAATCAGGCGCTCGCTCCCTATTCCTTCTATTCAGCCTTATTTCTT 2820
Db 2761 TGGTAACCTTTCAGAAATCAGGCGCTCGCTCCCTATTCCTTCTATTCAGCCTTATTTCTT 2820
QY 2821 ATGAGGAAGATCAGAGGCAAGGAGAGAACCTAGAAAAAATTTGTCAAGCCTATATGAAA 2880
Db 2821 ATGAGGAAGATCAGAGGCAAGGAGAGAACCTAGAAAAAATTTGTCAAGCCTATATGAAA 2880
QY 2881 CCAAACTTACTTTTGGAAAGTGCACATCATATGSCACCCACTAAAGATGAGTTGACT 2940
Db 2881 CCAAACTTACTTTTGGAAAGTGCACATCATATGSCACCCACTAAAGATGAGTTGACT 2940
QY 2941 GCAAGCCTGGGCTATTTCTGATGTGACCTGGAAAAAGATGTGCACCTCAGGCTGA 3000
Db 2941 GCAAGCCTGGGCTATTTCTGATGTGACCTGGAAAAAGATGTGCACCTCAGGCTGA 3000
QY 3001 TTGGACCCCTTCTGCTCTGCCACACTAACACACTGAACCCCTGCTCATGGGAGACAAGTGA 3060
Db 3001 TTGGACCCCTTCTGCTCTGCCACACTAACACACTGAACCCCTGCTCATGGGAGACAAGTGA 3060
QY 3061 CAGTACAGGAATTTGCTCTGTTTTCACCATCTTTGATGAGAACCAAAAGCTGTACTTCA 3120
Db 3061 CAGTACAGGAATTTGCTCTGTTTTCACCATCTTTGATGAGAACCAAAAGCTGTACTTCA 3120
QY 3121 CTGAAAAATATGAAAAAACTGCAGGGCTCCCTGCAATATCCAGATGGAAAGATCCCACTT 3180
Db 3121 CTGAAAAATATGAAAAAACTGCAGGGCTCCCTGCAATATCCAGATGGAAAGATCCCACTT 3180
QY 3181 TTTAAGAGAAATTAATGCTTCCATGCAATCAATGGCTACATAATGATACACTACCTGGCT 3240
Db 3181 TTTAAGAGAAATTAATGCTTCCATGCAATCAATGGCTACATAATGATACACTACCTGGCT 3240
QY 3241 TAGTAATGGCTCAGGATCAAAAGATTGATGCTGCTCAGCATGGGCAGCAATGAAA 3300
Db 3241 TAGTAATGGCTCAGGATCAAAAGATTGATGCTGCTCAGCATGGGCAGCAATGAAA 3300
QY 3301 ACATCCATTTCTATTCATTTCAAGTGCATGTGTCTACTGTACGAAAAAAGAGAGATATA 3360
Db 3301 ACATCCATTTCTATTCATTTCAAGTGCATGTGTCTACTGTACGAAAAAAGAGAGATATA 3360
QY 3361 AAATGGCACTGTACAACTCTATCCAGGTGTTTGTAGAGACAGTGAATGTTACCATCCA 3420
Db 3361 AAATGGCACTGTACAACTCTATCCAGGTGTTTGTAGAGACAGTGAATGTTACCATCCA 3420
QY 3421 AAGCTGGAATTTGGGGGGTGAATGCTTATTTGGGAGCATCTACATGCTGGGATGAGCA 3480
Db 3421 AAGCTGGAATTTGGGGGGTGAATGCTTATTTGGGAGCATCTACATGCTGGGATGAGCA 3480
QY 3481 CACTTTTCTGTGTACAGCAATAGTGTACAGACTCCCTGGGAATGGCTTCTGAGACA 3540
Db 3481 CACTTTTCTGTGTGTACAGCAATAGTGTACAGACTCCCTGGGAATGGCTTCTGAGACA 3540

Db 3481 CACTTTTCTGTGTGTACAGCAATAGTGTACAGACTCCCTGGGAATGGCTTCTGAGACA 3540
QY 3541 TTAGAGATTTTTCAGATTTACAGCTTCAGGACAAATATGACAGTGGGCCCCAAAGCTGGCCA 3600
Db 3541 TTAGAGATTTTTCAGATTTACAGCTTCAGGACAAATATGACAGTGGGCCCCAAAGCTGGCCA 3600
QY 3601 GACTTCATTTATCCGGATCAATCAATGCTGAGACCAAGAGGCCCTTTTCTTGGATCA 3660
Db 3601 GACTTCATTTATCCGGATCAATCAATGCTGAGACCAAGAGGCCCTTTTCTTGGATCA 3660
QY 3661 AGGTGGATCTGTGGCCACCAATGATTTATTCACGGCATCAAGACCAGGGTCCCTCAGA 3720
Db 3661 AGGTGGATCTGTGGCCACCAATGATTTATTCACGGCATCAAGACCAGGGTCCCTCAGA 3720
QY 3721 AGTTCCTCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGAGAAAGT 3780
Db 3721 AGTTCCTCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGAGAAAGT 3780
QY 3781 GGCAGACTTATCGAGAAATTCACCTGGAACCTTAATGCTCTTCTTGGCAATGTGATT 3840
Db 3781 GGCAGACTTATCGAGAAATTCACCTGGAACCTTAATGCTCTTCTTGGCAATGTGATT 3840
QY 3841 CATCTGGGATAAACACAATATTTTAAACCCCTCAATTTATGCTGTGATACATCCGTTTC 3900
Db 3841 CATCTGGGATAAACACAATATTTTAAACCCCTCAATTTATGCTGTGATACATCCGTTTC 3900
QY 3901 ACCCACTCATTTATGACATTCGAGCACTCTTGCATGAGATGAGTGGGCTGTGATTTAA 3960
Db 3901 ACCCACTCATTTATGACATTCGAGCACTCTTGCATGAGATGAGTGGGCTGTGATTTAA 3960
QY 3961 ATAGTTGCAGCATGCCATTTGGAAATGGAGAGATAAGCAATATCAGATGCACAGATTAAGT 4020
Db 3961 ATAGTTGCAGCATGCCATTTGGAAATGGAGAGATAAGCAATATCAGATGCACAGATTAAGT 4020
QY 4021 CTTCATCCTACTTTAACCAATATGTTTGGCACCTGGTCTCCTCAAAAAGCTGCATTCACC 4080
Db 4021 CTTCATCCTACTTTAACCAATATGTTTGGCACCTGGTCTCCTCAAAAAGCTGCATTCACC 4080
QY 4081 TCCAAGGGAGAGATTAATGCCCTGGAGACCTCAGTGTGATAATCCAAAAAGAGTGGCTGCAAG 4140
Db 4081 TCCAAGGGAGAGATTAATGCCCTGGAGACCTCAGTGTGATAATCCAAAAAGAGTGGCTGCAAG 4140
QY 4141 TGGACTTCCAGAAACAATGAAGTTCACAGAGATTAATCTCAGGGAGTAAATCTCTGC 4200
Db 4141 TGGACTTCCAGAAACAATGAAGTTCACAGAGATTAATCTCAGGGAGTAAATCTCTGC 4200
QY 4201 TTACCAAGCATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGA 4260
Db 4201 TTACCAAGCATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGA 4260
QY 4261 CTCTCTTTTTCAGAAATGGCAAAAGTAAAGGTTTTCAGGGAAATCAAGACTCCTTCACAC 4320
Db 4261 CTCTCTTTTTCAGAAATGGCAAAAGTAAAGGTTTTCAGGGAAATCAAGACTCCTTCACAC 4320
QY 4321 CTGTGGTGAACCTCTTAGACCCACCGTTACTGACTGCTACCTTCGAATTCACCCCCAGA 4380
Db 4321 CTGTGGTGAACCTCTTAGACCCACCGTTACTGACTGCTACCTTCGAATTCACCCCCAGA 4380
QY 4381 GTTGGGTGCACCAAGATTGCCCTGAGAGATGGAAGTCTGGGCTGCCAGAGCAGAGACTCT 4440
Db 4381 GTTGGGTGCACCAAGATTGCCCTGAGAGATGGAAGTCTGGGCTGCCAGAGCAGAGACTCT 4440
QY 4441 ACTGAGGGGTGGCCACTGCAGACCTGCCACTGCGGTCACTCTCCCTCTCAGCTCCAGG 4500
Db 4441 ACTGAGGGGTGGCCACTGCAGACCTGCCACTGCGGTCACTCTCCCTCTCAGCTCCAGG 4500
QY 4501 GCAGTGTCCCTCCCTGGCTTGCCTTCTACCTTTGTGCTAAATCTTAGCAGACACTGCTT 4560
Db 4501 GCAGTGTCCCTCCCTGGCTTGCCTTCTACCTTTGTGCTAAATCTTAGCAGACACTGCTT 4560
QY 4561 GAAGCCTCCTGAATTAATCATCATGCTCTGCAATTTCTTTGGTGGGGGGCCAGAGGGGT 4620
Db 4561 GAAGCCTCCTGAATTAATCATCATGCTCTGCAATTTCTTTGGTGGGGGGCCAGAGGGGT 4620

QY 4621 GCATCCAACTTAACCTTACTTCTTCTGACAGTCTCCAGATTACTCTTCC 4680
|||||
Db 4621 GCATCCAACTTAACCTTACTTCTTCTGACAGTCTCCAGATTACTCTTCC 4680
QY 4681 TTCCAAATTAAGTACGCAAGAGAGAGAGAAACCTGCATGAAGCATTTCTCCCTG 4740
|||||
Db 4681 TTCCAAATTAAGTACGCAAGAGAGAGAGAAACCTGCATGAAGCATTTCTCCCTG 4740
QY 4741 AAAAGTTAGGCTCTCAGAGTCACCACTTCTCTGTGTAGAAAAAAGTATGTGATGAAC 4800
|||||
Db 4741 AAAAGTTAGGCTCTCAGAGTCACCACTTCTCTGTGTAGAAAAAAGTATGTGATGAAC 4800
QY 4801 TTGAAAAAGATATTTATGATGTTGCGCGCCG 4832
|||||
Db 4801 TTGAAAAAGATATTTATGATGTTGCGCGCCG 4832

JUL 2
581
ID V19581 standard; cDNA; 4832 BP.
AC V19581;
DE 06-AUG-1998 (first entry)
KW Human factor VIII beta-domain deleted SQN deletion cDNA sequence.
KW Replication defective; recombinant retrovirus; RRV; therapeutic protein;
KW haemophilia; thrombosis; hypercoagulable disorder; liver disease; human;
KW hepatitis; thalassemia; phenylketonuria; Lesch-Nyhan syndrome; diabetes;
KW cystic fibrosis; Duchenne's Muscular Dystrophy; hypercholesterolemia;
KW hypopituitarism; adenine deaminase deficiency; HIV infection; anaemia;
KW Guacher's syndrome; high blood pressure; Alzheimer's disease; autoimmune;
KW inflammatory disease; factor VIII; ss.
OS Homo sapiens.
PN WO9800541-A2.
PD 08-JAN-1998.
PF 02-JUL-1997; U11784.
PR 04-JUN-1997; US-869309.
PR 03-JUL-1996; US-645601.
PR 13-AUG-1996; US-696381.
PA (CHIR) CHIRON CORP.
PI Allen JR, Barber JR, Boder M, Chang SMW, Chong K,
PI De LA VEGA D, Depolon J, Greengard J, Hsu DC, Ibanez CE,
PI Jolly DJ, Lee R, Mittelstaedt DM, Prussak CE, Respass JG;
PI WPI; 98-086966/08.
DR P-PSDB; W46246.
PT New replication defective recombinant retro-viruses - which can be
administered to provide long term systemic expression of therapeutic
protein in blood, useful in, e.g. treating hyper-coagulable
disorders

PS Example 28; Pages 210-213; 272pp; English.
CC This cDNA encodes the beta-domain deleted SQN deletion protein of human
CC factor VIII. This is used in the construction of recombinant retroviral
CC vectors expressing human factor VIII. The invention provides the
CC preparation of replication defective recombinant retrovirus (RRV)
CC expressing a therapeutic protein. The RRV preparation is resistant to
CC degradation by human complement and is capable of inducing long term
CC intravenously to a human. The long term systemic expression results in a
CC measurable level of the therapeutic protein being produced in the blood
CC of the human for a period of at least 30 days after the administration of
CC the RRV vector preparation. RRV's can be used for in vivo delivery of
CC therapeutic protein to treat, e.g. haemophilia A, haemophilia B,
CC thrombosis, hypercoagulable disorders, liver diseases such as hepatitis,
CC disorders such as thalassemia, phenylketonuria, Lesch-Nyhan syndrome,
CC severe combined immunodeficiency (SCID), cystic fibrosis, hypercholesterolemia,
CC Muscular Dystrophy, inherited emphysema, familial hypercholesterolemia,
CC diabetes, hypopituitarism, adenine deaminase deficiency, alpha1-
CC antitrypsin deficiency, Guacher's syndrome, anaemia, infections such as
CC HIV infection, high blood pressure, Alzheimer's disease, autoimmune or
CC inflammatory disease or graft versus host disease. RRV's are capable of
CC surviving inactivation in human serum thereby allowing efficient gene
CC transfer over prolonged periods of time.
SQ Sequence 4832 BP; 1381 A; 1089 C; 1055 G; 1307 T;

Query Match 100.0%; Score 4832; DB 1; Length 4832;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4832; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTCGAGCTAAAGATATTTAGAGAGAAATTAACCTTTTCTCTCCAGTTGAACATTTGT 60
Db 1 CTCGAGCTAAAGATATTTAGAGAGAAATTAACCTTTTCTCTCCAGTTGAACATTTGT 60
QY 61 AGCAATAAGTCATGCAATATAGAGCTCTCCACCTGCTCTTCTGCTTTGGATTCT 120
Db 61 AGCAATAAGTCATGCAATATAGAGCTCTCCACCTGCTCTTCTGCTTTGGATTCT 120
QY 121 GCTTAGTCCACAGAGATACTACTGCTGGTGCAGTGAAGTCTATGGACTATATGC 180
Db 121 GCTTAGTCCACAGAGATACTACTGCTGGTGCAGTGAAGTCTATGGACTATATGC 180
QY 181 AAAGTATCTCGGTGAGCTGCTGTGACGCAAGATTTCTCTAGAGTGCAAAATCTT 240
Db 181 AAAGTATCTCGGTGAGCTGCTGTGACGCAAGATTTCTCTAGAGTGCAAAATCTT 240
QY 241 TTCCATTCAACACCTCAGTCGTGTACAAAAAGACTGTTGTAGAAATTCAGGATCACC 300
Db 241 TTCCATTCAACACCTCAGTCGTGTACAAAAAGACTGTTGTAGAAATTCAGGATCACC 300
QY 301 TTTTCAACATCGCTAAGCCAAAGCCCTGTGATGGTCTGTAGGTCCTACATCCAGG 360
Db 301 TTTTCAACATCGCTAAGCCAAAGCCCTGTGATGGTCTGTAGGTCCTACATCCAGG 360
QY 361 CTGAGGTTATGATACAGTGTGATTAACCTTAAGAACATGGCTCCATCCTGTCAAGTC 420
Db 361 CTGAGGTTATGATACAGTGTGATTAACCTTAAGAACATGGCTCCATCCTGTCAAGTC 420
QY 421 TTTCATGCTGTGTGTATCTCTACTGAAAGCTTCTGAGGAGCTGAATATGATCAGA 480
Db 421 TTTCATGCTGTGTGTATCTCTACTGAAAGCTTCTGAGGAGCTGAATATGATCAGA 480
QY 481 CCAAGTCAAGGAGAGAAAGAGATTAAGTCTTCCCTGTGGAAGCCATACATATGTCT 540
Db 481 CCAAGTCAAGGAGAGAGAAAGAGATTAAGTCTTCCCTGTGGAAGCCATACATATGTCT 540
QY 541 GGCAGGCTCTGAAGAGAAATGTCATGAGCTTGCACCACTGTGCTTACTACTCAT 600
Db 541 GGCAGGCTCTGAAGAGAAATGTCATGAGCTTGCACCACTGTGCTTACTACTCAT 600
QY 601 ATCTTTCTCATGTGACCTGTGTAAGAAAGCTTGAATTCAGGCTCATTTGAGCCCTACTAG 660
Db 601 ATCTTTCTCATGTGACCTGTGTAAGAAAGCTTGAATTCAGGCTCATTTGAGCCCTACTAG 660
QY 661 TATGTAGAGAAAGGAGTCTGCGCAAGAAAGACACAGACCTTGACAAATTTACTACTAC 720
Db 661 TATGTAGAGAAAGGAGTCTGCGCAAGAAAGACACAGACCTTGACAAATTTACTACTAC 720
QY 721 TTTTGTCTGATTTGATGAAGGAGAAAGTGGCACTCAGAAACAAAGAACTCCTTGATGC 780
Db 721 TTTTGTCTGATTTGATGAAGGAGAAAGTGGCACTCAGAAACAAAGAACTCCTTGATGC 780
QY 781 AGGATAGGAGTGTGATCTGCTGCGGCTGAGCTTAAATGACACAGTCAATGTTATG 840
Db 781 AGGATAGGAGTGTGATCTGCTGCGGCTGAGCTTAAATGACACAGTCAATGTTATG 840
QY 841 TAAACAGGCTCTGCGCAAGTCTGATTTGATGCCACAGAAATCAGTCTATTTGGCATGTGA 900
Db 841 TAAACAGGCTCTGCGCAAGTCTGATTTGATGCCACAGAAATCAGTCTATTTGGCATGTGA 900
QY 901 TTGGAATGGGACCACTCTGGAAGTGCATCAATATTTCTGGAAGTGCACATTTCTTG 960
Db 901 TTGGAATGGGACCACTCTGGAAGTGCATCAATATTTCTGGAAGTGCACATTTCTTG 960
QY 961 TGAGGAACCATCGCCAGGCGTCTTGAATCTCGCCAAATTAATTTCTTACTGCTCAAA 1020
Db 961 TGAGGAACCATCGCCAGGCGTCTTGAATCTCGCCAAATTAATTTCTTACTGCTCAAA 1020

QY	1021	CACCTCTGATGACCTTGACAGTTTCTACTGTTTTGTTCATATCTCTCCACCAACATG	1080
Db	1021	CACCTTGTGATGACCTTGACAGTTTCTACTGTTTTGTTCATATCTCTCCACCAACATG	1080
QY	1081	ATGCGATGGAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGAAACCCCACTACGAATGA	1140
Db	1081	ATGCGATGGAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGAAACCCCACTACGAATGA	1140
QY	1141	AAAATTAATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAATGATGTGG	1200
Db	1141	AAAATTAATGAAGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAATGATGTGG	1200
QY	1201	TCAGGTTTGATGATGACAACTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGAAGC	1260
Db	1201	TCAGGTTTGATGATGACAACTCTCTCTTATCCAAATTCGCTCAGTTGCCAAGAAGC	1260
QY	1261	ATCCTAAACCTGGGTACATTACATTGCTGCTGAAGAAGAGGAACTGGGACTATGCTCCT	1320
Db	1261	ATCCTAAACCTGGGTACATTACATTGCTGCTGAAGAAGAGGAACTGGGACTATGCTCCT	1320
QY	1321	TAGTCTCTGCCCCCGATGACAGAGTTATAAAAGTCMAATATTTGAACAATGGCCCTCAGC	1380
Db	1321	TAGTCTCTGCCCCCGATGACAGAGTTATAAAAGTCMAATATTTGAACAATGGCCCTCAGC	1380
QY	1381	GGATTGGTAGGAAGTACAAAAAAGTCCGATTATGGCATACACAGATGAACCTTTAAGA	1440
Db	1381	GGATTGGTAGGAAGTACAAAAAAGTCCGATTATGGCATACACAGATGAACCTTTAAGA	1440
QY	1441	CTGCTGAAGCTATTCAGCATGAATCAGGAATCTGGGACCTTTACTTATGGGGAAGTTG	1500
Db	1441	CTGCTGAAGCTATTCAGCATGAATCAGGAATCTGGGACCTTTACTTATGGGGAAGTTG	1500
QY	1501	GAGACACACTGTGATTAATTAAGAATCAAGCAAGCAGACCATATAACATCTACCCCTC	1560
Db	1501	GAGACACACTGTGATTAATTAAGAATCAAGCAAGCAGACCATATAACATCTACCCCTC	1560
QY	1561	ACGGAATCACTGATGTCCGCTCTTGTATTCAAGGAGATTAACCAAAAGGTGTAAACATT	1620
Db	1561	ACGGAATCACTGATGTCCGCTCTTGTATTCAAGGAGATTAACCAAAAGGTGTAAACATT	1620
QY	1621	TGAAGGATTTTCCAATTCGCCAGGAGAAATATTCMAATATATAATGGACAGTGAAGTGA	1680
Db	1621	TGAAGGATTTTCCAATTCGCCAGGAGAAATATTCMAATATATAATGGACAGTGAAGTGA	1680
QY	1681	AAGATGGGCCAACTAAATCAGATCTCGGTGCTGACCCCGCTATTACTCTAGTTTGGTTA	1740
Db	1681	AAGATGGGCCAACTAAATCAGATCTCGGTGCTGACCCCGCTATTACTCTAGTTTGGTTA	1740
QY	1741	ATATGGAGAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGTCTCAAAAGAA	1800
Db	1741	ATATGGAGAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGTCTCAAAAGAA	1800
QY	1801	CTGTAGATCAAAAGAGAAACCAAGATTAATGTTCAGACCAAGAGAGATGTCAATCTGTTTCTG	1860
Db	1801	CTGTAGATCAAAAGAGAGAAACCAAGATTAATGTTCAGACCAAGAGAGATGTCAATCTGTTTCTG	1860
QY	1861	TATTTGATGAGAACCGAAGCTGTGTAACCTCACAGAGAAATATAACGCTTTCTCCCAATC	1920
Db	1861	TATTTGATGAGAACCGAAGCTGTGTAACCTCACAGAGAAATATAACGCTTTCTCCCAATC	1920
QY	1921	CAGCTGGAGTGACAGCTTGAGGATCCAGAGTTCCAAAGCCCTCAACATCATGACAGCATCA	1980
Db	1921	CAGCTGGAGTGACAGCTTGAGGATCCAGAGTTCCAAAGCCCTCAACATCATGACAGCATCA	1980
QY	1981	ATGCGTATGTTTTTGATAGTTTGAGGAGTTGTCAAGTTTGTGTTGATGAGGTGACATAGTGT	2040
Db	1981	ATGCGTATGTTTTTGATAGTTTGAGGAGTTGTCAAGTTTGTGTTGATGAGGTGACATAGTGT	2040
QY	2041	ACATTCTAAGCATTTGGAGCACAGACTGACTTCTTCTGTCTTCTTCTGAGATATACCT	2100
Db	2041	ACATTCTAAGCATTTGGAGCACAGACTGACTTCTTCTGTCTTCTTCTGAGATATACCT	2100
QY	2101	TCAAAACACAAAATGCTCTATGAAGACACACACTGACCCCTATTCCCAATTCCTCAGAGAAACTG	2160

Db	2101	TCAAACACAAAATG	GTCTATG	AAGACACACAC	ATCACCCCTAT	TCCCATTTCTC	GAGAGAAACTG	2160
QY	2161	TCCTCATGTCGAT	TGGAAAAACCC	AGTCTATG	GATTC	TGGGTGCCAC	ACTCAGACTTTC	2220
Db	2161	TCCTCATGTCGAT	TGGAAAAACCC	AGTCTATG	GATTC	TGGGTGCCAC	ACTCAGACTTTC	2220
QY	2221	GGAACAGAGGCAT	GCACCGCCCTT	ACTG	AAGGTTTCT	AGTTGTG	ACAAGAACACTG	2280
Db	2221	GGAACAGAGGCAT	GCACCGCCCTT	ACTG	AAGGTTTCT	AGTTGTG	ACAAGAACACTG	2280
QY	2281	ATTACGAGGACAGT	TATG	AAGATATTTC	AGCATACT	TGCTG	AGTAAAAACATGCC	2340
Db	2281	ATTACGAGGACAGT	TATG	AAGATATTTC	AGCATACT	TGCTG	AGTAAAAACATGCC	2340
QY	2341	AACCAAGAGAGCT	TCTCCAGAACCC	ACCAAGTCT	TGAACGCC	ATCAACGG	GAATAACTC	2400
Db	2341	AACCAAGAGAGCT	TCTCCAGAACCC	ACCAAGTCT	TGAACGCC	ATCAACGG	GAATAACTC	2400
QY	2401	GTACTACTCTTCA	GTCCAGATCA	GAGAAATG	CACTATG	ATACCAT	ATCAGTTGAA	2460
Db	2401	GTACTACTCTTCA	GTCCAGATCA	GAGAAATG	CACTATG	ATACCAT	ATCAGTTGAA	2460
QY	2461	TGAAGAAGGAAG	ATTTTGACAT	TTATGATG	AGGATG	AAATCAGAG	CCCCCGCAGCTTTC	2520
Db	2461	TGAAGAAGGAAG	ATTTTGACAT	TTATGATG	AGGATG	AAATCAGAG	CCCCCGCAGCTTTC	2520
QY	2521	AAAAAGAAACAC	GCACACTAT	TTTATG	CTG	CAGTGGAGAG	GCTCTGGGAT	2580
Db	2521	AAAAAGAAACAC	GCACACTAT	TTTATG	CTG	CAGTGGAGAG	GCTCTGGGAT	2580
QY	2581	GTAGCTCCCCAC	ATCTTCTA	GAACAA	CAGGGCTC	AGAGTGGC	AGTGTCCCTCAGTTC	2640
Db	2581	GTAGCTCCCCAC	ATCTTCTA	GAACAA	CAGGGCTC	AGAGTGGC	AGTGTCCCTCAGTTC	2640
QY	2641	AAGTGTTTTCC	AGCAATTTAC	TGATG	GCTCC	TTACTCAG	CCCTTATACCGTGGAG	2700
Db	2641	AAGTGTTTTCC	AGCAATTTAC	TGATG	GCTCC	TTACTCAG	CCCTTATACCGTGGAG	2700
QY	2701	TAAATGAACAT	TTGGGACTC	CTGGGCCAT	ATAT	AAGAGCAG	AAGTTGAAGATAAT	2760
Db	2701	TAAATGAACAT	TTGGGACTC	CTGGGCCAT	ATAT	AAGAGCAG	AAGTTGAAGATAAT	2760
QY	2761	TGGTACTTTCA	GAAATCAG	GGCTCTG	CTCCAT	TCTCAT	TCTAGCC	2820
Db	2761	TGGTACTTTCA	GAAATCAG	GGCTCTG	CTCCAT	TCTCAT	TCTAGCC	2820
QY	2821	ATGAGAGAATCA	GAGGCAAG	AGACAGAAC	CTAG	AAAACTTGT	CAAGCCTAATG	2880
Db	2821	ATGAGAGAATCA	GAGGCAAG	AGACAGAAC	CTAG	AAAACTTGT	CAAGCCTAATG	2880
QY	2881	CCAAACTTACT	TTTGGAAAG	TGCAACAT	CATATG	GCACCC	CTAAGAGT	2940
Db	2881	CCAAACTTACT	TTTGGAAAG	TGCAACAT	CATATG	GCACCC	CTAAGAGT	2940
QY	2941	GCAAAGCCTG	GGCTTATTCT	CTGATGT	GACCTG	GA	AAAAATGTGC	3000
Db	2941	GCAAAGCCTG	GGCTTATTCT	CTGATGT	GACCTG	GA	AAAAATGTGC	3000
QY	3001	TTGGACCCCT	TCTGTG	CTGCCAC	ACTA	ACAC	CTGAACCC	3060
Db	3001	TTGGACCCCT	TCTGTG	CTGCCAC	ACTA	ACAC	CTGAACCC	3060
QY	3061	CAGTACAGGA	ATTGCTCTG	TTTTCC	ACCATCT	TTGATG	AGACCAAAAG	3120
Db	3061	CAGTACAGGA	ATTGCTCTG	TTTTCC	ACCATCT	TTGATG	AGACCAAAAG	3120
QY	3121	CTGAATAAT	TGAAAGAA	ACTG	CAGGGCT	CCCTG	CAATATCC	3180
Db	3121	CTGAATAAT	TGAAAGAA	ACTG	CAGGGCT	CCCTG	CAATATCC	3180
QY	3181	TTAAAGGA	ATTATG	CGCTT	CCATG	CAATAG	GCTACAT	3240
Db	3181	TTAAAGGA	ATTATG	CGCTT	CCATG	CAATAG	GCTACAT	3240

```
Db 3181 TTAAAGAGAAATTATCGCTTCCATGCAATCAATGGGTACATAATGATACACTACTGGCT 3240
QY 3241 TAGTAATGGCTCAGATCAAAAGGATTGATGTAATCTGCTCAGCATGGGCAATGAAA 3300
Db 3241 TAGTAATGGCTCAGATCAAAAGGATTGATGTAATCTGCTCAGCATGGGCAATGAAA 3300
QY 3301 ACATCCATTTCTATTCATTTTCAAGTGCATGTGTTCACCTGTACGAAAAAAGAGAGTATA 3360
Db 3301 ACATCCATTTCTATTCATTTTCAAGTGCATGTGTTCACCTGTACGAAAAAAGAGAGTATA 3360
QY 3361 AAATGGCACTGTACAAATCTCTATTCAGAGGTGTTTGGAGACAGTGAATGTTACCATCCA 3420
Db 3361 AAATGGCACTGTACAAATCTCTATTCAGAGGTGTTTGGAGACAGTGAATGTTACCATCCA 3420
QY 3421 AAGCTGGAATTTGGGGGGTGAATGCTTATTTGGGAGACATCTACATGCTGGAGTGAACA 3480
Db 3421 AAGCTGGAATTTGGGGGGTGAATGCTTATTTGGGAGACATCTACATGCTGGAGTGAACA 3480
QY 3481 CACTTTTCTGTGTACAGCAATAGTGTACAGACTCCCTGGGAATGGCTTGTGACACA 3540
Db 3481 CACTTTTCTGTGTACAGCAATAGTGTACAGACTCCCTGGGAATGGCTTGTGACACA 3540
QY 3541 TTAGAGATTTTCAAGATTACAGCTTCAGGACAAATATGACAGTGGGCCCCAAAGCTGGCCA 3600
Db 3541 TTAGAGATTTTCAAGATTACAGCTTCAGGACAAATATGACAGTGGGCCCCAAAGCTGGCCA 3600
QY 3601 GACTTCATTTATTCGGGATCAATCAATGCTTGGAGACCAAGGAGCCCTTTCTTGATCA 3660
Db 3601 GACTTCATTTATTCGGGATCAATCAATGCTTGGAGACCAAGGAGCCCTTTCTTGATCA 3660
QY 3661 AGGTGATCTGTGGGACCAATATTTATTCACGGCATCAAGACCGGGTGGCCGTGAGA 3720
Db 3661 AGGTGATCTGTGGGACCAATATTTATTCACGGCATCAAGACCGGGTGGCCGTGAGA 3720
QY 3721 AGTTCTCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGAGAAAGT 3780
Db 3721 AGTTCTCAGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGAGAAAGT 3780
QY 3781 GGCAGACTTATCGAGGAAATTCACCTGGAACCTTAATGCTTCTTGGCAATGTGAT 3840
Db 3781 GGCAGACTTATCGAGGAAATTCACCTGGAACCTTAATGCTTCTTGGCAATGTGAT 3840
QY 3841 CATCTGGGATAAACACAATATTTTAAACCTCCCAATTTATGCTCGATACATCCGTTTGC 3900
Db 3841 CATCTGGGATAAACACAATATTTTAAACCTCCCAATTTATGCTCGATACATCCGTTTGC 3900
QY 3901 ACCCACTCATTTATAGCATTCGACACTCTTCGCATGGAGTTGATGGGCTGATTTAA 3960
Db 3901 ACCCACTCATTTATAGCATTCGACACTCTTCGCATGGAGTTGATGGGCTGATTTAA 3960
QY 3961 ATAGTGCAGATGCCATTGGGAATGGAGATGAAGCAATATCAGATGACAGATTACTG 4020
Db 3961 ATAGTGCAGATGCCATTGGGAATGGAGATGAAGCAATATCAGATGACAGATTACTG 4020
QY 4021 CTTCATCTTACTTTACCAATATGTTTGGCCACTGCTCTCTTCAAAAAGCTGCACTTACC 4080
Db 4021 CTTCATCTTACTTTACCAATATGTTTGGCCACTGCTCTCTTCAAAAAGCTGCACTTACC 4080
QY 4081 TCCAAGGGAGAGTAATGCTGAGACCTCAAGGTGAATATCCAAAAAGAGTGGCTGCAAG 4140
Db 4081 TCCAAGGGAGAGTAATGCTGAGACCTCAAGGTGAATATCCAAAAAGAGTGGCTGCAAG 4140
QY 4141 TGGACTTCCAGAAAGACAATGAAGTCAACAGAGTAATCTACTCAGGGAGTAAATCTCTGC 4200
Db 4141 TGGACTTCCAGAAAGACAATGAAGTCAACAGAGTAATCTACTCAGGGAGTAAATCTCTGC 4200
QY 4201 TTACCAAGCATGTATGTGAAGAGTTCCTCATCTCCAGACATCAAGATGGCATCAGTGA 4260
Db 4201 TTACCAAGCATGTATGTGAAGAGTTCCTCATCTCCAGACATCAAGATGGCATCAGTGA 4260
QY 4261 CTCTCTTTTTCAGAAATGGCAAAAGTAAAGTCTTTCAGGGAAATCAAGACTCCTTCACAC 4320
Db 4261 CTCTCTTTTTCAGAAATGGCAAAAGTAAAGTCTTTCAGGGAAATCAAGACTCCTTCACAC 4320
```

```
QY 4321 CTGTGTAAGTCTCTAGACCAACCGTTACTGACTCGCTACTCTGAAATTCACCCACAGA 4380
Db 4321 CTGTGTAAGTCTCTAGACCAACCGTTACTGACTCGCTACTCTGAAATTCACCCACAGA 4380
QY 4381 GTTGGGTGCACGATTGGCCCTGAGAGAGAGTCTTGCGCTGCGAGGACAGACTCT 4440
Db 4381 GTTGGGTGCACGATTGGCCCTGAGAGAGAGTCTTGCGCTGCGAGGACAGACTCT 4440
QY 4441 ACTGAGGTGCGCACTGCACACCTGCCACTGCCGTACTCTCCCTCTCAGCTCAGG 4500
Db 4441 ACTGAGGTGCGCACTGCACACCTGCCACTGCCGTACTCTCCCTCTCAGCTCAGG 4500
QY 4501 GCAGTGTCCCTCCCTGCTGCTGCTTCTACTTTGTGCTAAATCCTAGACACACTGCTT 4560
Db 4501 GCAGTGTCCCTCCCTGCTGCTGCTTCTACTTTGTGCTAAATCCTAGACACACTGCTT 4560
QY 4561 GAAAGCTCCTGAATTAATCATCATCAGTCCATTTCTTGTGGGGGCCAGAGGCT 4620
Db 4561 GAAAGCTCCTGAATTAATCATCATCAGTCCATTTCTTGTGGGGGCCAGAGGCT 4620
QY 4621 GCATCCAATTTAATTAATCTTACCTATTTCTGACGCTGCCAGATTAATCTCTTCC 4680
Db 4621 GCATCCAATTTAATTAATCTTACCTATTTCTGACGCTGCCAGATTAATCTCTTCC 4680
QY 4681 TTCCAATTAATTAAGCAAAAGAGTGAGAGAAACCTGCATGAAGCATTTCTTCCCTG 4740
Db 4681 TTCCAATTAATTAAGCAAAAGAGTGAGAGAAACCTGCATGAAGCATTTCTTCCCTG 4740
QY 4741 AAAAGTTAGCCTCTCAGAGTCCACACTTCTCTGTGTAGAAAAACATATGTGATGAAC 4800
Db 4741 AAAAGTTAGCCTCTCAGAGTCCACACTTCTCTGTGTAGAAAAACATATGTGATGAAC 4800
QY 4801 TTGAAAAAGATATTTATGATGTTGGGCGCGC 4832
Db 4801 TTGAAAAAGATATTTATGATGTTGGGCGCGC 4832

RESULT 3
ID T73164 standard; cDNA; 9354 BP.
AC T73164;
DE 08-APR-1998 (first entry)
KW Post-translational regulatory element; PRE; enhancer II; intronless gene;
KW surface antigen gene; cytoplasmic accumulation; targeted delivery;
KW near consensus splice sequence; blood coagulation factor; factor VIII;
OS Homo sapiens.
FH Key
FT CDS
FT 2965..7380
FT 5165..5174
FT misc_feature
FT /tag= a
FT /note= "5', near consensus site"
FT 5695..5703
FT misc_feature
FT /tag= c
FT /note= "5', near consensus site"
FT 6320..6328
FT misc_feature
FT /tag= d
FT /note= "5', near consensus site"
FT 6595..5603
FT misc_feature
FT /tag= e
FT /note= "5', near consensus site"
FT 7045..7053
FT misc_feature
FT /tag= f
FT /note= "5', near consensus site"
FT 7143..7152
FT misc_feature
FT /tag= g
FT /note= "5', near consensus site"
FT 3296..3312
FT misc_feature
FT /tag= h
FT /note= "3', near consensus site"
```

FT misc_feature 4798.4817
 FT /tag= 1
 FT /note= "3' near consensus site"
 FT misc_feature 5023.5045
 FT /tag= j
 FT /note= "3' near consensus site"
 FT misc_feature 5333.5355
 FT /tag= k
 FT /note= "3' near consensus site"
 FT misc_feature 5520.5538
 FT /tag= 1
 FT /note= "3' near consensus site"
 FT misc_feature 5604.5632
 FT /tag= m
 FT /note= "3' near consensus site"
 FT misc_feature 5717.5745
 FT /tag= n
 FT /note= "3' near consensus site"
 FT misc_feature 6239.6258
 FT /tag= o
 FT /note= "3' near consensus site"
 FT misc_feature 6658.6682
 FT /tag= p
 FT /note= "3' near consensus site"
 FT misc_feature 7159.7176
 FT /tag= q
 FT /note= "3' near consensus site"
 FT misc_feature 7196.7209
 FT /tag= r
 FT /note= "3' near consensus site"
 FT misc_feature 7289.7315
 FT /tag= s
 FT /note= "3' near consensus site"
 FT misc_feature 7411.7429
 FT /tag= t
 FT /note= "3' near consensus site"
 FT misc_feature 7611.8197
 FT /tag= u
 FT /note= "PRE sequence"
 FT /note= "PRE sequence"
 PN WO9733994-A1.
 PD 18-SEP-1997.
 PF 10-MAR-1997; U03561.
 PR 11-MAR-1996; US-683839.
 PA (IMMU-) IMMUNE RESPONSE CORP.
 PT Bidlingmaier S, Ill CR;
 DR WPI; 97-470874/43.
 DR P-PSDB; W23414.
 PT Vector for increased expression of intronless genes - comprises
 PT intronless gene with at least one near consensus splice sequence, a
 PT promoter and at least one viral cis-acting post-transcriptional
 PT regulatory element
 PS Example 1; Pages 21-31; 59pp; English.
 CC The present sequence represents human B-domain deleted factor VIII
 CC cDNA, and a post-translational regulatory element (PRE) of the
 CC Hepatitis B Virus, which is present 3' of the STOP codon for factor VIII.
 CC PRE sequences have been shown to function in cis to increase the
 CC steady-state levels of surface gene transcripts by facilitating
 CC cytoplasmic accumulation of these transcripts. The present sequence
 CC is part of a novel vector, comprising an intronless gene containing
 CC 1 or more near consensus splice sequences operably linked to a
 CC promoter sequence so that the gene is transcribed in a cell.
 CC Intronless gene transcripts which contain near consensus splice site
 CC sequences are believed to get tied up in the nucleus of the cell where
 CC splicing occurs, rather than being transported to the cytoplasm where
 CC they can be translated into proteins. The PRE sequences are transcribed
 CC along with the gene, causing export of the gene transcript from the
 CC nucleus into the cytoplasm of the cell. The vector can be used
 CC to increase the expression of an intronless gene containing at least one
 CC near consensus splice sites, preferably cDNA encoding a blood coagulation
 CC factor, particularly Factor VIII or IX. The complex allows the targeted
 CC delivery of the vector to a specific cell, e.g. hepatocytes when the
 CC ligand is an asialoglycoprotein which binds the asialoglycoprotein
 CC receptor present on their surface.

SQ Sequence 9354 BP; 2506 A; 2239 C; 2161 G; 2448 T;
 Query Match 94.6%; Score 4570.8; DB 1; Length 9354;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 4643; Conservative 0; Mismatches 12; Indels 45; Gaps 2;
 QY 14 TATTTAGAGAAGATTAACCTTTTGGCTTCACAGTGAACATTTGTAGCAATAG--T 70
 DB 2904 TATTTAGAGAAGATTAACCTTTTGGCTTCACAGTGAACATTTGTAGCAATAGCCAC 2963
 QY 71 CATGCAATAGAGCTCTCCACCTGCTCTTCTGTGCGCTTTGCGATTTGCTTAGTGC 130
 DB 2964 CATGTTATAGAGCTCTCCACCTGCTCTTCTGTGCGCTTTGCGATTTGCTTAGTGC 3023
 QY 131 CACCAGAAGATTAACCTGGTGACAGTGAACGTGATGGGACTATATGCAAGTATCT 190
 DB 3024 CACCAGAAGATTAACCTGGTGACAGTGAACGTGATGGGACTATATGCAAGTATCT 3083
 QY 191 CGGTGAGCTGCTGTGACCGCAAGATTCTCTCTAGAGTGGCCAAATCTTTCCATTCAA 250
 DB 3084 CGGTGAGCTGCTGTGACCGCAAGATTCTCTCTAGAGTGGCCAAATCTTTCCATTCAA 3143
 QY 251 CACCTCAGTCTGTACAAAAGACTCTGTTGTAGAAATTCACGGATCACCCTTTCAACAT 310
 DB 3144 CACCTCAGTCTGTACAAAAGACTCTGTTGTAGAAATTCACGGATCACCCTTTCAACAT 3203
 QY 311 CGCTAAGCCAAAGCCACCCTGGATGGTCTGCTAGGTCCTTACCATCCAGGCTGAGGTTTA 370
 DB 3204 CGCTAAGCCAAAGCCACCCTGGATGGTCTGCTAGGTCCTTACCATCCAGGCTGAGGTTTA 3263
 QY 371 TGATACAGTGGTCAATTACACTTAAGACATGAGCTTCCCATCTGTCAGTCTCATGCTGT 430
 DB 3264 TGATACAGTGGTCAATTACACTTAAGACATGAGCTTCCCATCTGTCAGTCTCATGCTGT 3323
 QY 431 TGGTGTATCTTACTGGAAGCTTCTGAGGAGCTGAATATGATGATCAGACCAAGTCAAG 490
 DB 3324 TGGTGTATCTTACTGGAAGCTTCTGAGGAGCTGAATATGATGATCAGACCAAGTCAAG 3383
 QY 491 GGAGAAAGAAATGATTAAGTCTTCCCTGGTGAAGCCATATATATGCTGGCAGGTCTCT 550
 DB 3384 GGAGAAAGAAATGATTAAGTCTTCCCTGGTGAAGCCATATATATGCTGGCAGGTCTCT 3443
 QY 551 GAAAGAGAAATGCTCAATGGCTCTGACCCACTGTGCTTACTACTCATATCTTTCTCA 610
 DB 3444 GAAAGAGAAATGCTCAATGGCTCTGACCCACTGTGCTTACTACTCATATCTTTCTCA 3503
 QY 611 TGTGACCTGTGAAAGACTTGAATTCAGGCTCATTTGAGCCCTACTAGTATGTAGAGA 670
 DB 3504 TGTGACCTGTGAAAGACTTGAATTCAGGCTCATTTGAGCCCTACTAGTATGTAGAGA 3563
 QY 671 AGGAGTCTGGCCAAAGAAAGACACAGACCTTGACAAATTTACTACTTTTGTCTGT 730
 DB 3564 AGGAGTCTGGCCAAAGAAAGACACAGACCTTGACAAATTTACTACTTTTGTCTGT 3623
 QY 731 ATTTGATGAAGGAAAGTTGGCACTCAGAAACAAGAACTCTTGATGACAGATAGGGA 790
 DB 3624 ATTTGATGAAGGAAAGTTGGCACTCAGAAACAAGAACTCTTGATGACAGATAGGGA 3683
 QY 791 TGTGCAATCTGCTGGGCTGCTGCTTAAATGACACAGTCAATGTTATGTAACAGGTC 850
 DB 3684 TGTGCAATCTGCTGGGCTGCTGCTTAAATGACACAGTCAATGTTATGTAACAGGTC 3743
 QY 851 TGTGCAAGTCTGATTTGATGACACAGAAATCAGTCTATTTGATGATGGAATGGG 910
 DB 3744 TGTGCAAGTCTGATTTGATGACACAGAAATCAGTCTATTTGATGATGGAATGGG 3803
 QY 911 CACCACTCTGAAGTGCACCTCAATATTCCTGAAGTGCACACATTTCTTGTGGAACCA 970
 DB 3804 CACCACTCTGAAGTGCACCTCAATATTCCTGAAGTGCACACATTTCTTGTGGAACCA 3863
 QY 971 TGCACAGGCTGCTTGAATCTGCAATATCTTCTTACTGCTCAACACTCTTGAT 1030
 DB 1030 TGCACAGGCTGCTTGAATCTGCAATATCTTCTTACTGCTCAACACTCTTGAT 1030

Db 3864 TCGCCAGGCGCTCCTTGGAAATCTGCCAATAACTTTCCTTACTGCTCAAAACACTCTTGAT 3923
QY 1031 GGACCTTGGACAGTTTTCTACTGTTTGTGCATATCTCTTCCCAACCAACATGATGGCATGA 1090
Db 3924 GGACCTTGGACAGTTTTCTACTGTTTGTGCATATCTCTTCCCAACCAACATGATGGCATGA 3983
QY 1091 AGCTTATGTCCAAGTAGACAGCTGTCCAGAGGAACCCCCAACTACGAATGAAAAATATGA 1150
Db 3984 AGCTTATGTCCAAGTAGACAGCTGTCCAGAGGAACCCCCAACTACGAATGAAAAATATGA 4043
QY 1151 AGAGCGGAGAGACTATGATGATGATCTTACTGATTTCTGAAATGGATGGGTGAGTTTGA 1210
Db 4044 AGAGCGGAGAGACTATGATGATGATCTTACTGATTTCTGAAATGGATGGGTGAGTTTGA 4103
QY 1211 TGATGACAACTCTCCTTCTTATCCAAATTCGCTCAGTTGCCAAGACATCCTTAAAC 1270
Db 4104 TGATGACAACTCTCCTTCTTATCCAAATTCGCTCAGTTGCCAAGACATCCTTAAAC 4163
QY 1271 TTGGGTACATTACATTGCTGCTGAAGAGGAGACTGGACTATGCTCCCTTAGTCTCGC 1330
Db 4164 TTGGGTACATTACATTGCTGCTGAAGAGGAGACTGGACTATGCTCCCTTAGTCTCGC 4223
QY 1331 CCCCAGTGCAGAGAGTTTAAAGTCAATATTTGAACAATGGCCCTCAGCGGATGGTAG 1390
Db 4224 CCCCAGTGCAGAGAGTTTAAAGTCAATATTTGAACAATGGCCCTCAGCGGATGGTAG 4283
QY 1391 GAAGTACAAAAAGTCCGATTATGGCATACACAGATGAACCTTTAAGACTCGTGAAGC 1450
Db 4284 GAAGTACAAAAAGTCCGATTATGGCATACACAGATGAACCTTTAAGACTCGTGAAGC 4343
QY 1451 TATTCAGCATGAATCAGGAATCTTGGGACCTTTACTTTATGGGAAATTGGAGACACACT 1510
Db 4344 TATTCAGCATGAATCAGGAATCTTGGGACCTTTACTTTATGGGAAATTGGAGACACACT 4403
QY 1511 GTTGATTATTTAAGAAATCAAGCAAGCAGACCATATTAACATCTACCTCAGCGAATCAC 1570
Db 4404 GTTGATTATTTAAGAAATCAAGCAAGCAGACCATATTAACATCTACCTCAGCGAATCAC 4463
QY 1571 TGATGTCCGCTCTTGTATTCAGAGAGATTACCAAAAGGTGTAAACATTTGAAGGATTT 1630
Db 4464 TGATGTCCGCTCTTGTATTCAGAGAGATTACCAAAAGGTGTAAACATTTGAAGGATTT 4523
QY 1631 TCCAATTTCTGCCAGAGAAATATTCAAATATAAATGGAAGTGAAGTGGGCC 1690
Db 4524 TCCAATTTCTGCCAGAGAAATATTCAAATATAAATGGAAGTGAAGTGGGCC 4583
QY 1691 AACTAAATCAGATCCTCGGTGCTGACCCCGCTATTACTCTAGTTTGTAAATATGAGAG 1750
Db 4584 AACTAAATCAGATCCTCGGTGCTGACCCCGCTATTACTCTAGTTTGTAAATATGAGAG 4643
QY 1751 AGATCTAGCTTAGGAGCTCATTTGCCCTCTCTCATCTGCTACAAAGAAATCTAGATCA 1810
Db 4644 AGATCTAGCTTAGGAGCTCATTTGCCCTCTCTCATCTGCTACAAAGAAATCTAGATCA 4703
QY 1811 AAGAGGAAACAGATATGTACAGCAAGAGAAATGTATCTGTTTCTGTATTTGATGA 1870
Db 4704 AAGAGGAAACAGATATGTACAGCAAGAGAAATGTATCTGTTTCTGTATTTGATGA 4763
QY 1871 GAACCGAAGCTGTACTCTACAGAGAAATATACACGCTTTCTCCCAATCCAGCTGGAGT 1930
Db 4764 GAACCGAAGCTGTACTCTACAGAGAAATATACACGCTTTCTCCCAATCCAGCTGGAGT 4823
QY 1931 GCAGCTTGAGAGTCCAGAGTTCCAGGCTTCAACATCATGACAGCATCAATGGCTATGT 1990
Db 4824 GCAGCTTGAGAGTCCAGAGTTCCAGGCTTCAACATCATGACAGCATCAATGGCTATGT 4883
QY 1991 TTTTGATAGTTTGCAGTTGTCAAGTTTGTTCAGATGAGGTGGCATTAAGTTCTAAG 2050
Db 4884 TTTTGATAGTTTGCAGTTGTGTCAAGTTTGTTCAGATGAGGTGGCATTAAGTTCTAAG 4943
QY 2051 CATTTGAGCAGAGACTGACTTCTTCTGTCTCTTCTCTGTGATATACCTTCAACACAA 2110
Db 4944 CATTTGAGCAGAGACTGACTTCTTCTGTCTCTTCTCTGTGATATACCTTCAACACAA 5003

QY 2111 AATGCTATGAAGACACACACTACCTATTTCCCATTTCTCAGAGAAACTGTCTCATGTC 2170
Db 5004 AATGCTATGAAGACACACACTACCTATTTCCCATTTCTCAGAGAAACTGTCTCATGTC 5063
QY 2171 GATGAAAAACCCAGGCTATAGATTTCTGGGGTCCACAACTCAGACTTTGGAACAGAGG 2230
Db 5064 GATGAAAAACCCAGGCTATAGATTTCTGGGGTCCACAACTCAGACTTTGGAACAGAGG 5123
QY 2231 CATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAAAGAACACTGGTATTTATCAGAGA 2290
Db 5124 CATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAAAGAACACTGGTATTTATCAGAGA 5183
QY 2291 CAGTTATGAAGATATTTCAAGCATCTTGTGAGTAAACAAATGCCATTTGAACCAAGAG 2350
Db 5184 CAGTTATGAAGATATTTCAAGCATCTTGTGAGTAAACAAATGCCATTTGAACCAAGAG 5243
QY 2351 CTCTCCAGAA-----CCCAACC 2368
Db 5244 CTCTCCAGAAATTCAGACACCTTAGCAGTAAAGCAATTTAATGCCACCCACC 5303
QY 2369 AGCTTTGAACCCCATCAACGGGAAATTAATCTGCTACTCTCAGTCAGATCAAGAGA 2428
Db 5304 AGCTTTGAACCCCATCAACGGGAAATTAATCTGCTACTCTCAGTCAGATCAAGAGA 5363
QY 2429 AATTGACTATGATATACCATATCAGTTGAATGAAGAGAAAGATTTGACATTTATGA 2488
Db 5364 AATTGACTATGATATACCATATCAGTTGAATGAAGAGAAAGATTTGACATTTATGA 5423
QY 2489 TGAGGATGAAATCAGAGCCCCGACGCTTCAAAAGAAACACGACATATTTATTTGC 2548
Db 5424 TGAGGATGAAATCAGAGCCCCGACGCTTCAAAAGAAACACGACATATTTATTTGC 5483
QY 2549 TGCAGTGAGAGGCTCTGGGATTTAGGATAGTAGTCCCCACATGTTCTAAGAAACAG 2608
Db 5484 TGCAGTGAGAGGCTCTGGGATTTAGGATAGTAGTCCCCACATGTTCTAAGAAACAG 5543
QY 2609 GGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGAAATTTACTGATGG 2668
Db 5544 GGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGAAATTTACTGATGG 5603
QY 2669 CTCCTTACTCAGCCCTTATACCGTGGAGAACTAAATGAACATTTGGGACTCCTGGGGCC 2728
Db 5604 CTCCTTACTCAGCCCTTATACCGTGGAGAACTAAATGAACATTTGGGACTCCTGGGGCC 5663
QY 2729 ATATATAAGCAGAGATGGAAGTAAATATCATGTGTAACCTTTCAGAAATCAGGCCCTCG 2788
Db 5664 ATATATAAGCAGAGATGGAAGTAAATATCATGTGTAACCTTTCAGAAATCAGGCCCTCG 5723
QY 2789 TCCCTATTCCTTCTATTTAGCCTTATTTCTTATGAGAGAGATCAGAGCAAGAGCAGA 2848
Db 5724 TCCCTATTCCTTCTATTTAGCCTTATTTCTTATGAGAGAGATCAGAGCAAGAGCAGA 5783
QY 2849 ACCTAGAAAAAATTTGTCAAGCCTAATGAACCAAACTTACTTTTGGAAAGTGCAACA 2908
Db 5784 ACCTAGAAAAAATTTGTCAAGCCTAATGAACCAAACTTACTTTTGGAAAGTGCAACA 5843
QY 2909 TCATATGGCACCACATAAGATGAGTTTGAAGTGAAGCCCTGGGCTTATTTCTGTGATGT 2968
Db 5844 TCATATGGCACCACATAAGATGAGTTTGAAGTGAAGCCCTGGGCTTATTTCTGTGATGT 5903
QY 2969 TGACCTGAAAAAAGATGTCAAGCCTGATTTGACCCCTTCTGTGCTGCCACACTAA 3028
Db 5904 TGACCTGAAAAAAGATGTCAAGCCTGATTTGACCCCTTCTGTGCTGCCACACTAA 5963
QY 3029 CACACTGAACCTGTCTATGGAGAGAGTGAACAGTACAGGAATTTGCTGTTTTTCAC 3088
Db 5964 CACACTGAACCTGTCTATGGAGAGAGTGAACAGTACAGGAATTTGCTGTTTTTCAC 6023
QY 3089 CATCTTGTATGAGACCAAAAGCTGTGTAATTTCACTGAAATATGAAAGAACTGCAGGGC 3148
Db 6024 CATCTTGTATGAGACCAAAAGCTGTGTAATTTCACTGAAATATGAAAGAACTGCAGGGC 6083

```
QY 3149 TCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAAATATCGCTTCCATGCAAT 3208
Db 6084 TCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAAATATCGCTTCCATGCAAT 6143
QY 3209 CAATGGCTACATATGATACACTGCTTGGCTTAGTAATGGCTCAGAGTCAAGAGATTGCG 3268
Db 6144 CAATGGCTACATATGATACACTGCTTGGCTTAGTAATGGCTCAGAGTCAAGAGATTGCG 6203
QY 3269 ATGGATCTGCTCAGCATGGGAGCAATGAAACATCTTATTCATTTTCAGTGACA 3328
Db 6204 ATGGATCTGCTCAGCATGGGAGCAATGAAACATCTTATTCATTTTCAGTGACA 6263
QY 3329 TGTGTTCACTGTACGAAAAAAGAGAGTATAAATGGCACTGTACAACTCTATCCAGG 3388
Db 6264 TGTGTTCACTGTACGAAAAAAGAGAGTATAAATGGCACTGTACAACTCTATCCAGG 6323
QY 3389 TGTGTTGAGACAGTGAATGTATACCATCCAAAGCTGGAATTGGCGGGTGAATGCTT 3448
Db 6324 TGTGTTGAGACAGTGAATGTATACCATCCAAAGCTGGAATTGGCGGGTGAATGCTT 6383
QY 3449 TATTTGGGAGCATCTACATGCTGGGATGAGCACACATTTTCTGTTGTAACGAAATAGTG 3508
Db 6384 TATTTGGGAGCATCTACATGCTGGGATGAGCACACATTTTCTGTTGTAACGAAATAGTG 6443
QY 3509 TCAGACTCCCTGGGAATGGCTTCTGAGACACATTTAGAGATTTTCAGATTACAGCTCAGG 3568
Db 6444 TCAGACTCCCTGGGAATGGCTTCTGAGACACATTTAGAGATTTTCAGATTACAGCTCAGG 6503
QY 3569 ACAATATGACAGTGGGCCCCAAAGCTGGCCAGACTTATTCGGGATCAATCAATGC 3628
Db 6504 ACAATATGACAGTGGGCCCCAAAGCTGGCCAGACTTATTCGGGATCAATCAATGC 6563
QY 3629 CTGAGCACCAAGAGCCCTTTCTTGATCAAGTGGATCTGTTGGCCCAATGATAT 3688
Db 6564 CTGAGCACCAAGAGCCCTTTCTTGATCAAGTGGATCTGTTGGCCCAATGATAT 6623
QY 3689 TCAGGGCATCAAGACCCAGGGTCCGTCAGAAAGTCTCCAGCCTTACATCTCTCAGTT 3748
Db 6624 TCAGGGCATCAAGACCCAGGGTCCGTCAGAAAGTCTCCAGCCTTACATCTCTCAGTT 6683
QY 3749 TATCATCATGTATAGTCTTGATGGAGAGAGTGGCAGACTTATCGAGAAATTCACATGG 3808
Db 6684 TATCATCATGTATAGTCTTGATGGAGAGAGTGGCAGACTTATCGAGAAATTCACATGG 6743
QY 3809 AACCTTAATGCTCTCTTTGGCAATGGATTCAATCGGGATAAACACAATATTTTAA 3868
Db 6744 AACCTTAATGCTCTCTTTGGCAATGGATTCAATCGGGATAAACACAATATTTTAA 6803
QY 3869 CCTTCCAAATATATGCTCGATACATCCGTTTGCACCAACTCATTTATGCAATTCGACAGAC 3928
Db 6804 CCTTCCAAATATATGCTCGATACATCCGTTTGCACCAACTCATTTATGCAATTCGACAGAC 6863
QY 3929 TCTTCGATGAGTGTGAGTGGCTGTGATTTAAATAGTTGCAGCATGCCATTTGGAATGA 3988
Db 6864 TCTTCGATGAGTGTGAGTGGCTGTGATTTAAATAGTTGCAGCATGCCATTTGGAATGA 6923
QY 3989 GAGTAAAGCAATATCAGATGCACAGATTACTGCTTCAATCTTACTTTACCAATATGTTGC 4048
Db 6924 GAGTAAAGCAATATCAGATGCACAGATTACTGCTTCAATCTTACTTTACCAATATGTTGC 6983
QY 4049 CACCTGCTCTCTTCAAAAGCTCGACTTCACCTCCAAGGAGAGTAATGCCGTGAGACC 4108
Db 6984 CACCTGCTCTCTTCAAAAGCTCGACTTCACCTCCAAGGAGAGTAATGCCGTGAGACC 7043
QY 4109 TCAGTGAATATCCAAAAGAGTGGCTGCAAGTGGACTTCAGAGACAATGAAAGTCAAC 4168
Db 7044 TCAGTGAATATCCAAAAGAGTGGCTGCAAGTGGACTTCAGAGACAATGAAAGTCAAC 7103
QY 4169 AGGAGTAATCTACTCAGGAGTAAATATCTGCTTACCAAGCATGTATGGAAGAGTTCTCT 4228
Db 7104 AGGAGTAATCTACTCAGGAGTAAATATCTGCTTACCAAGCATGTATGGAAGAGTTCTCT 7163
QY 4229 CATCTCAAGCAGTCAAGATGGCCATGACTGACTCTTTTTCAGAAATGGCAAGTAATA 4288
```

```
Db 7164 CATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTCTTTTTCAGAAATGGCAAGTAATA 7223
QY 4289 GGTTTTTCAGGGAATCAAGACTCTTCACACCTGTGTGTAAGTCTCTAGACCAACCGTT 4348
Db 7224 GGTTTTTCAGGGAATCAAGACTCTTCACACCTGTGTGTAAGTCTCTAGACCAACCGTT 7283
QY 4349 ACTGACTCGCTACTTCCGAATTTCACCCCCAGAGTTGGGTGCACCAAGATTGCCCTGAGGAT 4408
Db 7284 ACTGACTCGCTACTTCCGAATTTCACCCCCAGAGTTGGGTGCACCAAGATTGCCCTGAGGAT 7343
QY 4409 GGAGTCTGCGGCTGCCAGGACACAGACCTCTTACTGAGGGTGGCCACTGCAGCAGCTGCC 4468
Db 7344 GGAGTCTGCGGCTGCCAGGACACAGACCTCTTACTGAGGGTGGCCACTGCAGCAGCTGCC 7403
QY 4469 ACTGCCGTCACCTCTCCCTCCTCAGCTCCAGGCGAGTGTCCCTCCCTGGCTTCTTA 4528
Db 7404 ACTGCCGTCACCTCTCCCTCCTCAGCTCCAGGCGAGTGTCCCTCCCTGGCTTCTTA 7463
QY 4529 CCTTTGTCTAAATCTTAGCAGACACTGCCCTTGAAGCCTCCTGAATTAATCATCAGT 4588
Db 7464 CCTTTGTCTAAATCTTAGCAGACACTGCCCTTGAAGCCTCCTGAATTAATCATCAGT 7523
QY 4589 CCTGCATTTCTTTGGTGGGGGCCAGAGGGTGCATCCAAATTAATTAATCTTAATCTTA 4648
Db 7524 CCTGCATTTCTTTGGTGGGGGCCAGAGGGTGCATCCAAATTAATTAATCTTAATCTTA 7583
QY 4649 TTTTCTCAGCTGCTCCAG 4668
Db 7584 TTTTCTCAGGCGATCTCAG 7603
```

```
RESULT 4
Q76016
ID 076016 standard; cdna; 4629 BP.
AC 076016;
DT 20-JUL-1995 (first entry)
DE B-domain deleted Factor-VIII.
KW Factor-VIII; blood-clotting; hemophilia A; gene therapy;
KW adenovirus; vector; ss.
OS Homo sapiens.
PN WO9429471-A.
PD 22-DEC-1994.
PF 13-APR-1994; U04075.
PR 10-JUN-1993; US-074920.
PR 25-MAR-1994; US-218335.
PA (GENE-) GENETIC THERAPY INC.
PI Connelly S, Kaleko M, Smith T;
DR WPI; 95-036495/05.
DR P-PSDB; R67709.
PT New adenoviral vectors for treatment of haemophilia - contg. a
PT DNA sequence encoding a clotting factor, partic. Factor VIII or
PT Factor IX
PS Disclosure; Fig. 17A-17C; 116pp; English.
CC Human Factor-VIII cDNA, from which the B domain had been deleted, was
CC used to construct recombinant adenovirus vectors that produced
CC therapeutic levels of the clotting factor when administered to an
CC animal host, potentially providing hemophilia A gene therapy.
SQ Sequence 4629 BP; 1318 A; 1051 C; 1018 G; 1242 T;
```

```
Query Match 92.6%; Score 4474.2; DB 1; Length 4629;
Best Local Similarity 98.3%; Pred. No. 0;
Matches 4549; Conservative 0; Mismatches 38; Indels 42; Gaps 1;
QY 72 ATGCAATAGAGCTCTCCACCTGCTTCTTCTGTGCTTTTGGCAATTCGTTAGTGCC 131
Db 1 ATGCAATAGAGCTCTCCACCTGCTTCTTCTGTGCTTTTGGCAATTCGTTAGTGCC 60
QY 132 ACCAGAGATACTACTGGTGCAGTGAAGTGTCAATGGGACTATATGCAAGTGTCTC 191
Db 61 ACCAGAGATACTACTGGTGCAGTGAAGTGTCAATGGGACTATATGCAAGTGTCTC 120
```

QY 192 GGTGAGCTGCTGTGGAGCGAAGATTTCTCCTAGAGTGCCAAATCTTTCCATTCAAC 251
|||||
Db 121 GGTGAGCTGCTGTGGAGCGAAGATTTCTCCTAGAGTGCCAAATCTTTCCATTGAAC 180
QY 252 ACCTCAGTCGTGTACAAAAAGACTCTGTTGTGAATTCAGGATCACCCTTTCAACATC 311
|||||
Db 181 ACCTCAGTCGTGTACAAAAAGACTCTGTTGTGAATTCAGGATCACCCTTTCAACATC 240
QY 312 GCTAAGCCAAAGCCACCCTGATGGCTGTGCTAGAGTCCATCCATCCAGGCTGAGTTAT 371
|||||
Db 241 GCTAAGCCAAAGCCACCCTGATGGCTGTGCTAGAGTCCATCCATCCAGGCTGAGTTAT 300
QY 372 GATACAGTGTCTATTACACTTAAGAACATGGCTCCCATCCTGTGAGTCTTCACTGTG 431
|||||
Db 301 GATACAGTGTCTATTACACTTAAGAACATGGCTCCCATCCTGTGAGTCTTCACTGTG 360
QY 432 GGTGTATCTACTGGAAGCTTCTGAGGGAGCTGAATATGATGATCAGACCAAGTCAAGG 491
|||||
Db 361 GGTGTATCTACTGGAAGCTTCTGAGGGAGCTGAATATGATGATCAGACCAAGTCAAGG 420
QY 492 GAGAAAGAAAGATGATAAAGCTTCCCTGTGGAAGCCATACATATGTCTGAGAGTCTG 551
|||||
Db 421 GAGAAAGAAAGATGATAAAGCTTCCCTGTGGAAGCCATACATATGTCTGAGAGTCTG 480
QY 552 AAAGAGATGGTCCAAATGGCTCTGACCCACTGTGCTTACTACTCATATCTTCTCAT 611
|||||
Db 481 AAAGAGATGGTCCAAATGGCTCTGACCCACTGTGCTTACTACTCATATCTTCTCAT 540
QY 612 GTGAGCTGGTAAAGACTTGAATTCAGGCTCATTTGAGCCCTACTAGTATGTAGAA 671
|||||
Db 541 GTGAGCTGGTAAAGACTTGAATTCAGGCTCATTTGAGCCCTACTAGTATGTAGAA 600
QY 672 GGGAGTCTGGCCAAAGAAAGACACAGACCTTGACAAATTTATACTACTTTTGTCTGA 731
|||||
Db 601 GGGAGTCTGGCCAAAGAAAGACACAGACCTTGACAAATTTATACTACTTTTGTCTGA 660
QY 732 TTGTATGAAAGGAAAAAGTTGGCACTCAGAAACAAAGAACTCTTGATGAGGATAGGAT 791
|||||
Db 661 TTGTATGAAAGGAAAAAGTTGGCACTCAGAAACAAAGAACTCTTGATGAGGATAGGAT 720
QY 792 GCTGATCTGTGCTGGGCTTGCCCTAAATATGCACACAGTCAATGTTATGAAACAGTCT 851
|||||
Db 721 GCTGATCTGTGCTGGGCTTGCCCTAAATATGCACACAGTCAATGTTATGAAACAGTCT 780
QY 852 CTGCCAGGTCTGATTTGATGCCACAGAAATCAGTCTATTTGGCATGTGATGGAATGGC 911
|||||
Db 781 CTGCCAGGTCTGATTTGATGCCACAGAAATCAGTCTATTTGGCATGTGATGGAATGGC 840
QY 912 ACCACTCTGAAGTGCACTCAATATTCCTCGAAGTGACACACATTTCTGTGAGAACCAT 971
|||||
Db 841 ACCACTCTGAAGTGCACTCAATATTCCTCGAAGTGACACACATTTCTGTGAGAACCAT 900
QY 972 CGCCAGGCGTCTTGAAATCTGCGCAATTAACCTTCTACTGCTCAAAACACTCTTGATG 1031
|||||
Db 901 CGCCAGGCGTCTTGAAATCTGCGCAATTAACCTTCTACTGCTCAAAACACTCTTGATG 960
QY 1032 GACCTTGGACAGTTTCTACTGTTTGTCAATCTCTCCCAACCAACATGATGCGATGAA 1091
|||||
Db 961 GACCTTGGACAGTTTCTACTGTTTGTCAATCTCTCCCAACCAACATGATGCGATGAA 1020
QY 1092 GCTTATGTCAAAGTAGACAGCTGTCCAGAGAAACCCCAACTAGCAATGAAAAATATGAA 1151
|||||
Db 1021 GCTTATGTCAAAGTAGACAGCTGTCCAGAGAAACCCCAACTAGCAATGAAAAATATGAA 1080
QY 1152 GAAGCGAAGACTATGATGATGATCTTACTGATTTGAAATGATGATGGTTCAGTTGAT 1211
|||||
Db 1081 GAAGCGAAGACTATGATGATGATCTTACTGATTTGAAATGATGATGGTTCAGTTGAT 1140
QY 1212 GATGACAACTCTCTCTCTTATCCAAATTCGTCAGTTGGCAAGAGCATCCTAAACT 1271
|||||
Db 1141 GATGACAACTCTCTCTCTTATCCAAATTCGTCAGTTGGCAAGAGCATCCTAAACT 1200
QY 1272 TGGTACATTACATGCTGCTGAAGAGAGAGACTGGGACTATGCTCCCTTAGTCTCGCC 1331

|||||
Db 1201 TGGTACATTACATGCTGCTGAAGAGAGAGACTGGACTATGCTCCCTTAGTCTCGCC 1260
QY 1332 CCCGATGACAGAGATTATAAAAGTCAATATTTGAACAATGGCCCTCAGCGGATGGTAGG 1391
|||||
Db 1261 CCCGATGACAGAGATTATAAAAGTCAATATTTGAACAATGGCCCTCAGCGGATGGTAGG 1320
QY 1392 AAGTACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGTGAAGCT 1451
|||||
Db 1321 AAGTACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGTGAAGCT 1380
QY 1452 ATTACATGAATCAGAAATCTTGGACCTTTACTTTATGGGAAGTTGGAGACACTG 1511
|||||
Db 1381 ATTACATGAATCAGAAATCTTGGACCTTTACTTTATGGGAAGTTGGAGACACTG 1440
QY 1512 TTGATTATATTTAAGATCAAGCAAGCAGACCATATTAACATCTACCTCAGGAATCACT 1571
|||||
Db 1441 TTGATTATATTTAAGATCAAGCAAGCAGACCATATTAACATCTACCTCAGGAATCACT 1500
QY 1572 GATGTCGCTCTTGTATTTCAAGGAGATTTACCAAAAGGTGTAAACATTTGAAGATTTT 1631
|||||
Db 1501 GATGTCGCTCTTGTATTTCAAGGAGATTTACCAAAAGGTGTAAACATTTGAAGATTTT 1560
QY 1632 CCAATTTCTGCCAGAGAAATATTTCAATATTAATGAACAGTGAAGTGAAGAGGCCCA 1691
|||||
Db 1561 CCAATTTCTGCCAGAGAAATATTTCAATATTAATGAACAGTGAAGTGAAGAGGCCCA 1620
QY 1692 ACTAAATCAGATCTCGGTGCTGACCCGCTATTTACTCTAGTTTGTATATGAGAGAGA 1751
|||||
Db 1621 ACTAAATCAGATCTCGGTGCTGACCCCTATTTACTCTAGTTTGTATATGAGAGAGA 1680
QY 1752 GATCTAGCTTCAAGACTCATTTGGCCCTCTCTCATCTGCTACAAAGAAATCTGTGATCAA 1811
|||||
Db 1681 GATCTAGCTTCAAGACTCATTTGGCCCTCTCTCATCTGCTACAAAGAAATTTGTGATCAA 1740
QY 1812 AGAGAAAAACAGATATGTACAGCAAGAGAAATGTATCTCTGTTTCTGTATTTGATGAG 1871
|||||
Db 1741 AGAGAAAAACAGATATGTACAGCAAGAGAAATGTATCTCTGTTTCTGTATTTGATGAG 1800
QY 1872 AACGGAAGCTGTACTCTCACAGAGAAATATACAAGCTTTCTCCCAATCCAGCTGAGAGTG 1931
|||||
Db 1801 AACGGAAGCTGTACTCTCACAGAGAAATATACAAGCTTTCTCCCAATCCAGCTGAGAGTG 1860
QY 1932 CAGCTTGAGGATCAGAGATTTCAGAGCTTCCAAACATCATGACACAGCATCAATGGCTATGTT 1991
|||||
Db 1861 CAGCTTGAGGATCAGAGATTTCAGAGCTTCCAAACATCATGACACAGCATCAATGGCTATGTT 1920
QY 1992 TTGTATAGTTTGAGTGTGACGTTTGTGATGAGGTGGCATACTGTGATCTTCTAAGC 2051
|||||
Db 1921 TTGTATAGTTTGAGTGTGACGTTTGTGATGAGGTGGCATACTGTGATCTTCTAAGC 1980
QY 2052 ATTGAGACACAGACTGACTTCTTCTGTCTTCTCTCTGTGATATACCTTCAAAACACAAA 2111
|||||
Db 1981 ATTGAGACACAGACTGACTTCTTCTGTCTTCTCTCTGTGATATACCTTCAAAACACAAA 2040
QY 2112 ATGCTATGAAGACACACACTCAACCTATTTCCATTTCTCAGAGAGAAACTGTTCATGTG 2171
|||||
Db 2041 ATGCTATGAAGACACACACTCAACCTATTTCCATTTCTCAGAGAGAAACTGTTCATGTG 2100
QY 2172 ATGGAACAAACCCAGGCTATGATTTCTGGGGTCCACAACCTGAGACTTTCGGAACAGAGGC 2231
|||||
Db 2101 ATGGAACAAACCCAGGCTATGATTTCTGGGGTCCACAACCTGAGACTTTCGGAACAGAGGC 2160
QY 2232 ATGACCGCTTACTGAAGTTTCTAGTTGTGACAAAGAACACTGGTATTTATTCAGAGAC 2291
|||||
Db 2161 ATGACCGCTTACTGAAGTTTCTAGTTGTGACAAAGAACACTGGTATTTATTCAGAGAC 2220
QY 2292 AGTTATGAAGATTTTACAGCAATCTTGTGAGTAAAAACAAATGCCATTTGAACCAAGAGC 2351
|||||
Db 2221 AGTTATGAAGATTTTACAGCAATCTTGTGAGTAAAAACAAATGCCATTTGAACCAAGAGC 2280
QY 2352 TTCTCCAGAA-----CCACCA 2369
|||||

Db 2281 TTCTCCAGAATTCAGAGACCCTAGCACTAGGCCAAAAGCAATTTAATGCCACCCACCA 2340
QY 2370 GTCTTGAAACGCCATCAACGGGAAATACTGCTACTCTTCAGTCAGATCAAGAGAA 2429
Db 2341 GTCTTGAAACGCCATCAACGGGAAATACTGCTACTCTTCAGTCAGATCAAGAGAA 2400
QY 2430 ATGACTATGATGATACCATATCAGTGAATGAAGAAGAAATTTGACATTATGAT 2489
Db 2401 ATTGACTATGATGATACCATATCAGTGAATGAAGAAGAAATTTGACATTATGAT 2460
QY 2490 GAGATGAAAAATCAGAGCCCCCGCAGCTTTCAAAAAGAAAACAGCACATATTTATGCT 2549
Db 2461 GAGATGAAAAATCAGAGCCCCCGCAGCTTTCAAAAAGAAAACAGCACATATTTATGCT 2520
QY 2550 GCAGTGAAGAGGCTCTGGGATTATGGGATGAGTAGCTCCCCACATGTTCTAAGAAAAG 2609
Db 2521 GCAGTGAAGAGGCTCTGGGATTATGGGATGAGTAGCTCCCCACATGTTCTAAGAAAAG 2580
QY 2610 GCTCAGAGTGGCAGTGTCCCTCAGTTCAGAAAGTGTTCAGAGAAATTTACTGATGC 2669
Db 2581 GCTCAGAGTGGCAGTGTCCCTCAGTTCAGAAAGTGTTCAGAGAAATTTACTGATGC 2640
QY 2670 TCCCTTACTCAGCCCTTATACCGTGAAGAACTAAATGACATTTGGGACTCTGGGCCA 2729
Db 2641 TCCCTTACTCAGCCCTTATACCGTGAAGAACTAAATGACATTTGGGACTCTGGGCCA 2700
QY 2730 TATATAAGAGCAGAAGTTGAAGATAATATCATGTGTAACCTTCAGAAATCAGGCCCTGCT 2789
Db 2701 TATATAAGAGCAGAAGTTGAAGATAATATCATGTGTAACCTTCAGAAATCAGGCCCTGCT 2760
QY 2790 CCCATTCCTTCTATTTAGCCCTTATTTCTTATGAGAGAGATCAGAGCAAGAGCAGAA 2849
Db 2761 CCCATTCCTTCTATTTAGCCCTTATTTCTTATGAGAGAGATCAGAGCAAGAGCAGAA 2820
QY 2850 CCTAGAAAAAAGCTTTGTCAAGCCTAATGAACAACAAACTTACTTTTGAAGTGCAACAT 2909
Db 2821 CCTAGAAAAAAGCTTTGTCAAGCCTAATGAACAACAAACTTACTTTTGAAGTGCAACAT 2880
QY 2910 CATATGGCACCCACTAAGATGAGTTTGACTGCAAAAGCCTGGGCTTATTTCTCTGATGT 2969
Db 2881 CATATGGCACCCACTAAGATGAGTTTGACTGCAAAAGCCTGGGCTTATTTCTCTGATGT 2940
QY 2970 GACCTGGA AAAAGATGTGCACCTCAGGCTGATTTGGACCCCTTCTGCTGCCACACTAAC 3029
Db 2941 GACCTGGA AAAAGATGTGCACCTCAGGCTGATTTGGACCCCTTCTGCTGCCACACTAAC 3000
QY 3030 ACACGTGAACCCCTGCTCATGGGAGACACAGTACAGTACAGGAATTTGCTCTGTTTTCACC 3089
Db 3001 ACACGTGAACCCCTGCTCATGGGAGACACAGTACAGTACAGGAATTTGCTCTGCGAGGCT 3060
QY 3090 ATCTTTGATGAGACCAAAAGCTGGTACTTTCACCTGAAAATATGGAAGAAAAGTGAGGCT 3149
Db 3061 ATCTTTGATGAGACCAAAAGCTGGTACTTTCACCTGAAAATATGGAAGAAAAGTGAGGCT 3120
QY 3150 CCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAAATATCGCTTCCATGCAATC 3209
Db 3121 CCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAAATATCGCTTCCATGCAATC 3180
QY 3210 AATGGCTACATAATGATACACTACCTGGCTTAGTAATGGCTCAGGATCAAAAGATTGCA 3269
Db 3181 AATGGCTACATAATGATACACTACCTGGCTTAGTAATGGCTCAGGATCAAAAGATTGCA 3240
QY 3270 TGGTATCTGCTCAGCATGGGAGCAATGAAAACATCCATTTCTATTCTATTCAGTGACAT 3329
Db 3241 TGGTATCTGCTCAGCATGGGAGCAATGAAAACATCCATTTCTATTCTATTCAGTGACAT 3300
QY 3330 GTGTTACTGTACGAAAAAAGAGAGATTAATAATGGCACTGTACAATCTCTATCCAGGT 3389
Db 3301 GTGTTACTGTACGAAAAAAGAGAGATTAATAATGGCACTGTACAATCTCTATCCAGGT 3360
QY 3390 GTTTTGGACAGAGTGAAGATTTACCATCCAAAAGCTGGAATTTGGCGGGTGAATGCCCT 3449
Db 3361 GTTTTGGACAGAGTGAAGATTTACCATCCAAAAGCTGGAATTTGGCGGGTGAATGCCCT 3420

QY 3450 ATTGGCAGCATCTACATGCTGGGATGAGCACACACTTTTCTGCTGTACAGCAATAAGTCT 3509
Db 3421 ATTGGCAGCATCTACATGCTGGGATGAGCACACACTTTTCTGCTGTACAGCAATAAGTCT 3480
QY 3510 CAGACTCCCTGGGAATGGCTCTCTGACACATTAGAGATTTTCAGATTTACAGCTTCAGCA 3569
Db 3481 CAGACTCCCTGGGAATGGCTCTCTGACACATTAGAGATTTTCAGATTTACAGCTTCAGCA 3540
QY 3570 CAATATGACAGTGGGCCCCCAAGCTGGCCAGACTTCATTTATCCGGATCAATCAATGCC 3629
Db 3541 CAATATGACAGTGGGCCCCCAAGCTGGCCAGACTTCATTTATCCGGATCAATCAATGCC 3600
QY 3630 TGGAGCACCAGAGAGCCCTTTCTTGATCAAGGTGATCTGTGGCACCATGATTTAT 3689
Db 3601 TGGAGCACCAGAGAGCCCTTTCTTGATCAAGGTGATCTGTGGCACCATGATTTAT 3660
QY 3690 CACGGCATCAAGACCAGGGGTGCCCGTCAGAAAGTTCACAGCCCTCTACATCTCTCAGTTT 3749
Db 3661 CACGGCATCAAGACCAGGGGTGCCCGTCAGAAAGTTCACAGCCCTCTACATCTCTCAGTTT 3720
QY 3750 ATCATCATGTATAGTCTTGATGGGAAGAGTGGCAGACTTATCGAGGAAATTCACATGGA 3809
Db 3721 ATCATCATGTATAGTCTTGATGGGAAGAGTGGCAGACTTATCGAGGAAATTCACATGGA 3780
QY 3810 ACCTTAATGGTCTTCTTTGGCAATGTGATTCATCTGGGATTAACAATATTTTAAAC 3869
Db 3781 ACCTTAATGGTCTTCTTTGGCAATGTGATTCATCTGGGATTAACAATATTTTAAAC 3840
QY 3870 CCTCAATTTATTTGCTGATACATCCGTTTGCACCCCAACTCATTTAGCATTCGACGACT 3929
Db 3841 CCTCAATTTATTTGCTGATACATCCGTTTGCACCCCAACTCATTTAGCATTCGACGACT 3900
QY 3930 CTTCGATGGAGTTGATGGGCTGTGATTTAATAGTTGCAGCATGCCATTGGGAATGAG 3989
Db 3901 CTTCGATGGAGTTGATGGGCTGTGATTTAATAGTTGCAGCATGCCATTGGGAATGAG 3960
QY 3990 AGTAAGCAATATCAGATGACAGATTACTGCTTCACTCACTTACCAATATGTTTGC 4049
Db 3961 AGTAAGCAATATCAGATGACAGATTACTGCTTCACTCACTTACCAATATGTTTGC 4020
QY 4050 ACCGTGCTCTCTTCAAAAAGCTGCACCTCACTCCAAAGGAGAGTAATGCCTGAGACT 4109
Db 4021 ACCGTGCTCTCTTCAAAAAGCTGCACCTCACTCCAAAGGAGAGTAATGCCTGAGACT 4080
QY 4110 CAGTGATTAATCCAAAAGAGTGGCTGCAAGTGAGACTTCAGAAAGCAATGAAGTGACA 4169
Db 4081 CAGTGATTAATCCAAAAGAGTGGCTGCAAGTGAGACTTCAGAAAGCAATGAAGTGACA 4140
QY 4170 GGAGTAACCTACTCAGGAGTAAATCTCTGCTTACACGATGTATGGAAGAGTTCTC 4229
Db 4141 GGAGTAACCTACTCAGGAGTAAATCTCTGCTTACACGATGTATGGAAGAGTTCTC 4200
QY 4230 ATCTCAGCAGTCAAGATGGCCATCAGTGGACTCTTTTTTCAGAAATGGCAAGTAAG 4289
Db 4201 ATCTCAGCAGTCAAGATGGCCATCAGTGGACTCTTTTTTCAGAAATGGCAAGTAAG 4260
QY 4290 GTTTTTCAGGGAATCAAGACTCCTTCACACCTGTGTGAACCTCTAGAACCCACCGTTA 4349
Db 4261 GTTTTTCAGGGAATCAAGACTCCTTCACACCTGTGTGAACCTCTAGAACCCACCGTTA 4320
QY 4350 CTGACTGCTACCTTCGAATTCACCCCAAGAGTTGGGTGCACCAAGATTGCCCTGAGATG 4409
Db 4321 CTGACTGCTACCTTCGAATTCACCCCAAGAGTTGGGTGCACCAAGATTGCCCTGAGATG 4380
QY 4410 GAGGTCTGGGCTGCGAGGCACAGAGCTCTACTGAGGTTGGCAGCTGACACCTGCCA 4469
Db 4381 GAGGTCTGGGCTGCGAGGCACAGAGCTCTACTGAGGTTGGCAGCTGACACCTGCCA 4440
QY 4470 CTGCCGTACCTCTCCCTCCTCAGCTCCAGGGCAGTGTCCCTGCCCTTGCCTTCTAC 4529
Db 4441 CTGCCGTACCTCTCCCTCCTCAGCTCCAGGGCAGTGTCCCTGCCCTTGCCTTCTAC 4500

QY 4530 CTTTGCTAAATCCTAGACAGACTGCCCTTGAGCCTCCTGAATTAATCATCATGATC 4589
|||||
Db 4501 CTTTGCTAAATCCTAGACAGACTGCCCTTGAGCCTCCTGAATTAATCATCATGATC 4560
QY 4590 CTGCATTCTTTGGTGGGGGCGCAGGAGGGTGCATCCAATTAACTTAACCTTACCTAT 4649
|||||
Db 4561 CTGCATTCTTTGGTGGGGGCGCAGGAGGGTGCATCCAATTAACTTAACCTTACCTAT 4620
QY 4650 TTTCTGCAG 4658
|||||
Db 4621 TTTCTGCAG 4629

RESULT 5

V23339
ID V23339 standard; DNA; 4670 BP.
AC V23339;
17-AUG-1998 (first entry)
Human Factor-VIII gene lacking central B domain.
Factor-VIII; blood clotting; human; synthetic gene; codon usage; ss.
OS Homo sapiens.
OS Synthetic.
PN WO9812207-A1.
PD 26-MAR-1998.
PF 18-SEP-1997; U16639.
PR 20-SEP-1996; US-717294.
PA (GEHO) GEN HOSPITAL CORP.
PI Haas J, Seed B;
DR WPI: 98-217200/19.
PT New synthetic eukaryotic gene(s) - in which non-preferred or less
PT preferred codon(s) are replaced to provide high level expression in
PT mammalian cell(s)
PS Example 3; Fig 12; 92pp; English.
CC This gene codes for a human Factor-VIII protein that lacks the
CC central B domain (amino acids 760-1639) of the native protein. In
CC a novel, claimed synthetic gene (see V23288), non-preferred or
CC less preferred codons of the native gene are replaced by codons
CC favored by highly expressed human genes to provide high-level
CC expression in mammalian cells. The synthetic gene was assembled
CC from 29 pairs of oligonucleotides (see V23340-97) which served as
CC PCR templates. Synthetic genes of the invention (see also
CC V23289-91) are used for production of recombinant proteins in
CC mammalian cells at levels of at least 500% of those obtained using
CC the natural genes. They can also be used in gene therapy.
SQ Sequence 4670 BP; 1325 A; 1072 C; 1036 G; 1237 T;

very Match 92.3%; Score 4462; DB 1; Length 4670;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 4520; Conservative 0; Mismatches 10; Indels 42; Gaps 1;

QY 129 GCCACCAGAGATACTACTGGGTGACGTGGAAGTGTATATGCAAAAGTAT 188
|||||
Db 84 GCCACCAGAGATACTACTGGGTGACGTGGAAGTGTATATGCAAAAGTAT 143
QY 189 CTCGGTAGCTGCTGTGACGCAAGATTCTCTAGAGTCCAAATCTTTTCCATTC 248
|||||
Db 144 CTCGGTAGCTGCTGTGACGCAAGATTCTCTAGAGTCCAAATCTTTTCCATTC 203
QY 249 AACACCTCAGTCTGTACAAAAGACTCTGTTGTAGAAATTCAGGATCAGCTTTTCAAC 308
|||||
Db 204 AACACCTCAGTCTGTACAAAAGACTCTGTTGTAGAAATTCAGGATCAGCTTTTCAAC 263
QY 309 ATCGCTAAGCCCAAGGCCACCTGGATGGGTCTGTAGTCTTACCATCCAGGCTGAGGTT 368
|||||
Db 264 ATCGCTAAGCCCAAGGCCACCTGGATGGGTCTGTAGTCTTACCATCCAGGCTGAGGTT 323
QY 369 TATGATACAGTGTGATTAACCTTAAGAAATGGCTTCCCATCTGTCACTTTCATGCT 428
|||||
Db 324 TATGATACAGTGTGATTAACCTTAAGAAATGGCTTCCCATCTGTCACTTTCATGCT 383
QY 429 GTTGGTATCTTACTGAAAGCTTCTGAGGAGCTGAATATGATGATGAGACAGTCAA 488
|||||

Db 384 GTTGGTATCTTACTGAAAGCTTCTGAGGGAGCTGAATATGATGATCAGACAGTCAA 443
QY 489 AGGAGAAGAGATGATTAAGTCTTCCCTGGTGAAGCCATACATATGCTGGCAGTTC 548
|||||
Db 444 AGGAGAAGAGATGATTAAGTCTTCCCTGGTGAAGCCATACATATGCTGGCAGTTC 503
QY 549 CTGAAGAAGATGTCCATAGGCCCTCTGACCCACTGTGCCTTACCTACATATCTTCT 608
|||||
Db 504 CTGAAGAAGATGTCCATAGGCCCTCTGACCCACTGTGCCTTACCTACATATCTTCT 563
QY 609 CATGTGACCTGGTAAAGACTTGAATTCAGGCCCTCATTTGAGCCCTACTAGTATGTA 668
|||||
Db 564 CATGTGACCTGGTAAAGACTTGAATTCAGGCCCTCATTTGAGCCCTACTAGTATGTA 623
QY 669 GAAGGAGTCTGGCCCAAGGAAAGACACAGACCTTGCACAATTTATCTACTTTTGTCT 728
|||||
Db 624 GAAGGAGTCTGGCCCAAGGAAAGACACAGACCTTGCACAATTTATCTACTTTTGTCT 683
QY 729 GTATTTGATGAAGGAAAGTTGGCACTCAGAAACAAGAACTCCTTGATGCAGATAGG 788
|||||
Db 684 GTATTTGATGAAGGAAAGTTGGCACTCAGAAACAAGAACTCCTTGATGCAGATAGG 743
QY 789 GATGCTGCATCTGCTCGGGCCTGGCCTAAATGCACACAGTCAATGGTTATGTAACAGG 848
|||||
Db 744 GATGCTGCATCTGCTCGGGCCTGGCCTAAATGCACACAGTCAATGGTTATGTAACAGG 803
QY 849 TCTCTGCAGGTCTGATTTGATGGACACAGGAATCAGTCTATGGCATGTGATGGAATG 908
|||||
Db 804 TCTCTGCAGGTCTGATTTGATGGACACAGGAATCAGTCTATGGCATGTGATGGAATG 863
QY 909 GGCACCACTCCTGAAGTGCACCTCAATATCTCTGAGAGTTCACACATTTCTTGAGAGAAC 968
|||||
Db 864 GGCACCACTCCTGAAGTGCACCTCAATATCTCTGAGAGTTCACACATTTCTTGAGAGAAC 923
QY 969 CATGCCAGGCGCTCTTGGAAATCTGCCCAATACTTCTTACTGCTCAACACACTTGTG 1028
|||||
Db 924 CATGCCAGGCGCTCTTGGAAATCTGCCCAATACTTCTTACTGCTCAACACACTTGTG 983
QY 1029 ATGACCTTGGACAGTTTCTACTGTTTGTGATATCTCTTCCACCAACATGATGGCATG 1088
|||||
Db 984 ATGACCTTGGACAGTTTCTACTGTTTGTGATATCTCTTCCACCAACATGATGGCATG 1043
QY 1089 GAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGAGAACCCCACTACGAATGAAATAAT 1148
|||||
Db 1044 GAAGCTTATGTCAAAGTAGACAGCTGTCCAGAGAGAACCCCACTACGAATGAAATAAT 1103
QY 1149 GAAGAAGCGGAGAGATATGATGATCTTACTGATCTTGAATGATGTGTCAGGTTT 1208
|||||
Db 1104 GAAGAAGCGGAGAGATATGATGATCTTACTGATCTTGAATGATGTGTCAGGTTT 1163
QY 1209 GATGATGACAACTCTCTCTTCTTATCCAAATTCGCTCAGTTGCCAAGAAGATCCTAAA 1268
|||||
Db 1164 GATGATGACAACTCTCTCTTCTTATCCAAATTCGCTCAGTTGCCAAGAAGATCCTAAA 1223
QY 1269 ACTTGGTACATTACATTTGCTGCTGAAGAGAGAGACTGGGACTATGCTCCCTTAGTCTC 1328
|||||
Db 1224 ACTTGGTACATTACATTTGCTGCTGAAGAGAGAGACTGGGACTATGCTCCCTTAGTCTC 1283
QY 1329 GCCCCGATGACAGAGTATTAAGTCAATATTTGAACAATGGCCCTCAGCGGATTTGGT 1388
|||||
Db 1284 GCCCCGATGACAGAGTATTAAGTCAATATTTGAACAATGGCCCTCAGCGGATTTGGT 1343
QY 1389 AGGAAGTACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGTGA 1448
|||||
Db 1344 AGGAAGTACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTTAAGACTCGTGA 1403
QY 1449 GCTATTCAGCATGAATCAGGAATCTTGGGACCTTACTTTATGGGGAAGTTGGAGACACA 1508
|||||
Db 1404 GCTATTCAGCATGAATCAGGAATCTTGGGACCTTACTTTATGGGGAAGTTGGAGACACA 1463
QY 1509 CTGTGATTTATTTAAGATCAAGCAAGCAGACCATATTAACATCTACCTCAGGGAATC 1568
|||||
Db 1464 CTGTGATTTATTTAAGATCAAGCAAGCAGACCATATTAACATCTACCTCAGGGAATC 1523

OY 1569 ACTGATGTCGCTCTTGTATTCAGAGATATACCAAAAGGTGTAAACATTTGAAGAT 1628
|||||
Db 1524 ACTGATGTCGCTCTTGTATTCAGAGATATACCAAAAGGTGTAAACATTTGAAGAT 1583
OY 1629 TTTCCAATTCGCCAGAGAAATATTCAAATATAATGACAGTGAAGATGGG 1688
|||||
Db 1584 TTTCCAATTCGCCAGAGAAATATTCAAATATAATGACAGTGAAGATGGG 1643
OY 1689 CCAACTAAATCAGATCCCTCGGTGCTGACCCGCTATTACTCTAGTTTTCGTTAATGAG 1748
|||||
Db 1644 CCAACTAAATCAGATCCCTCGGTGCTGACCCGCTATTACTCTAGTTTTCGTTAATGAG 1703
OY 1749 AGAGATCTAGCTTCAGAGACTCAATTGGCCCTCTCCTCATCTGCTACAAGAATCTAGAT 1808
|||||
Db 1704 AGAGATCTAGCTTCAGAGACTCAATTGGCCCTCTCCTCATCTGCTACAAGAATCTAGAT 1763
OY 1809 CAAAGAGAAACAGATATATGTCAGACAAGAGAAATGTCATCCTGTTTCTGTATTGAT 1868
|||||
Db 1764 CAAAGAGAAACAGATATATGTCAGACAAGAGAAATGTCATCCTGTTTCTGTATTGAT 1823
OY 1869 GAGAACCGAAGCTGTACTCCTCAGAGAAATATACAACGCTTCTCCCAATCCAGCTGGA 1928
|||||
Db 1824 GAGAACCGAAGCTGTACTCCTCAGAGAAATATACAACGCTTCTCCCAATCCAGCTGGA 1883
OY 1929 GTGCAGCTTGAGGATCCAGAGTTCCAAAGCCTCCAACATCATGACAGCATCAATGGCTAT 1988
|||||
Db 1884 GTGCAGCTTGAGGATCCAGAGTTCCAAAGCCTCCAACATCATGACAGCATCAATGGCTAT 1943
OY 1989 GTTTTGTAGATTGTGCAGTTGTGCAGTTGTGTTCGATGAGGTGGCATACTGGTACATTCTA 2048
|||||
Db 1944 GTTTTGTAGATTGTGCAGTTGTGTTCGATGAGGTGGCATACTGGTACATTCTA 2003
OY 2049 AGCATTTGAGACAGACTGACTTCTTCTGTCTTCTCTGATATACCTTCAAAAC 2108
|||||
Db 2004 AGCATTTGAGACAGACTGACTTCTTCTGTCTTCTCTGATATACCTTCAAAAC 2063
OY 2109 AAAATGCTCTATGAAGACACACTCAACCTATTCCCATTTCTCAGAGAAACTGTCTTCATG 2168
|||||
Db 2064 AAAATGCTCTATGAAGACACACTCAACCTATTCCCATTTCTCAGAGAAACTGTCTTCATG 2123
OY 2169 TCGATGGAAAAACCCAGGCTATGAGATTCTGGGGTCCACAACTGAGACTTTGGAAACAGA 2228
|||||
Db 2124 TCGATGGAAAAACCCAGGCTATGAGATTCTGGGGTCCACAACTGAGACTTTGGAAACAGA 2183
OY 2229 GGCATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAAGAACACTGGTATTATTACGAG 2288
|||||
Db 2184 GGCATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAAGAACACTGGTATTATTACGAG 2243
OY 2289 GACAGTTATGAAGATATTTACAGACTTCTGCTGAGTAAAAACAATGCCATTGAACCAAGA 2348
|||||
Db 2244 GACAGTTATGAAGATATTTACAGACTTCTGCTGAGTAAAAACAATGCCATTGAACCAAGA 2303
OY 2349 AGCTTCTCCAGAA-----CCCA 2366
|||||
Db 2304 AGCTTCTCCAGAAATTCAGACACCCCTAGCAGTAAAGCAAAAGCAATTTAATGCCACCCCA 2363
OY 2367 CCAGTCTTGAAGCCCATCAACGGGAAATTAACCTGTAATCTTCACTCAGTCAAGAGAG 2426
|||||
Db 2364 CCAGTCTTGAAGCCCATCAACGGGAAATTAACCTGTAATCTTCACTCAGTCAAGAGAG 2423
OY 2427 GAAATTCATATGATATACCATATCAGTTGAATGAAGAAGATTTTGACATTAT 2486
|||||
Db 2424 GAAATTCATATGATATACCATATCAGTTGAATGAAGAAGATTTTGACATTAT 2483
OY 2487 GATGAGATGAAAAATCAGAGCCCGCAGCTTTCAAAAAGAAACAGACACTATTTTAT 2546
|||||
Db 2484 GATGAGATGAAAAATCAGAGCCCGCAGCTTTCAAAAAGAAACAGACACTATTTTAT 2543
OY 2547 GCTGAGTGAGAGGCTCTGGGATATGAGATGAGTACCTCCACATGTTCTAAGAAAC 2606
|||||
Db 2544 GCTGAGTGAGAGGCTCTGGGATATGAGATGAGTACCTCCACATGTTCTAAGAAAC 2603

OY 2607 AGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGGAATTTACTGAT 2666
|||||
Db 2604 AGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTCCAGGAATTTACTGAT 2663
OY 2667 GGCTCCTTTACTCAGCCCTTATACCGGTGAGAACTAAATGAACATTTGGGACTCTGGGG 2726
|||||
Db 2664 GGCTCCTTTACTCAGCCCTTATACCGGTGAGAACTAAATGAACATTTGGGACTCTGGGG 2723
OY 2727 CCATATATAAGACAGAAAGTTGAAGATAAATATCATGTGTAACCTTTCAGAAATCAGCCCTCT 2786
|||||
Db 2724 CCATATATAAGACAGAAAGTTGAAGATAAATATCATGTGTAACCTTTCAGAAATCAGCCCTCT 2783
OY 2787 GGTCCCTATTCCTTCTATTTAGCCCTTATTTCTATGAGGAAGATCAGAGCAAGAGCA 2846
|||||
Db 2784 GGTCCCTATTCCTTCTATTTAGCCCTTATTTCTTATGAGGAAGATCAGAGCAAGAGCA 2843
OY 2847 GAACCTAGAAAAAATTGTCAAGCCTAATGAACCAAAACTTACTTTTGAAGTGCAA 2906
|||||
Db 2844 GAACCTAGAAAAAATTGTCAAGCCTAATGAACCAAAACTTACTTTTGAAGTGCAA 2903
OY 2907 CATCATATGGCACCACCTAAAGATGAGTTTGACTGCAAAAGCCTGGGCTTATTTCTGTAT 2966
|||||
Db 2904 CATCATATGGCACCACCTAAAGATGAGTTTGACTGCAAAAGCCTGGGCTTATTTCTGTAT 2963
OY 2967 GTTGACCTGGAAGAAAGATGTGCACTCAGGCCCTGATTTGACCCCTTCTGTGTCACACACT 3026
|||||
Db 2964 GTTGACCTGGAAGAAAGATGTGCACTCAGGCCCTGATTTGACCCCTTCTGTGTCACACACT 3023
OY 3027 AACACACTGAACCCCTGCTCATGGGAGACACAAGTGACAGTACAGAAATTTGCTGTCTTTC 3086
|||||
Db 3024 AACACACTGAACCCCTGCTCATGGGAGACACAAGTGACAGTACAGAAATTTGCTGTCTTTC 3083
OY 3087 ACCATCTTTGATGAGACCAAAAGCTGTACTTCACTGAAAAATATGGAAGAACTGCAGG 3146
|||||
Db 3084 ACCATCTTTGATGAGACCAAAAGCTGTACTTCACTGAAAAATATGGAAGAACTGCAGG 3143
OY 3147 GCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAATTATCGCTTCCATGCA 3206
|||||
Db 3144 GCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAATTATCGCTTCCATGCA 3203
OY 3207 ATCAATGGCTACATAATGATATCACTACCTGGCTTAGTAATGGCTCAGATCAAGGAT 3266
|||||
Db 3204 ATCAATGGCTACATAATGATATCACTACCTGGCTTAGTAATGGCTCAGATCAAGGAT 3263
OY 3267 CGATGGTATCTGCTCAGCATGGGCAGCAATGAACATCCATTCTATTCAATTTCAGTGA 3326
|||||
Db 3264 CGATGGTATCTGCTCAGCATGGGCAGCAATGAACATCCATTCTATTCAATTTCAGTGA 3323
OY 3327 CATGTGTCACTGTACGAAAAAAGAGAGATTAATGCGCACTGTACAATCTCTATCCA 3386
|||||
Db 3324 CATGTGTCACTGTACGAAAAAAGAGAGATTAATGCGCACTGTACAATCTCTATCCA 3383
OY 3387 GGTGTTTTGAGACAGTGAATGTTTACCATTCAAAAGCTGGAATTTGGCGGGTGAATGC 3446
|||||
Db 3384 GGTGTTTTGAGACAGTGAATGTTTACCATTCAAAAGCTGGAATTTGGCGGGTGAATGC 3443
OY 3447 CTTATTGGCGAGCATCTACATGCTGGGATGAGCACACTTTTCTGCTGTACAGCAATAAG 3506
|||||
Db 3444 CTTATTGGCGAGCATCTACATGCTGGGATGAGCACACTTTTCTGCTGTACAGCAATAAG 3503
OY 3507 TGTGAGACTCCCTGGGAATGCTTCTGGACACATTAGAGATTTTCAGATTTACAGCTTCA 3566
|||||
Db 3504 TGTGAGACTCCCTGGGAATGCTTCTGGACACATTAGAGATTTTCAGATTTACAGCTTCA 3563
OY 3567 GGACAATATGACAGTGGGCCCAAGCTGGCGAGACTTCAATTAATCCGGATCAATCAAT 3626
|||||
Db 3564 GGACAATATGACAGTGGGCCCAAGCTGGCGAGACTTCAATTAATCCGGATCAATCAAT 3623
OY 3627 GCCTGAGCACCAAGAGCCCTTTTCTTGATCAAGGTGATCTGTTGGACCAATGAT 3686
|||||
Db 3624 GCCTGAGCACCAAGAGCCCTTTTCTTGATCAAGGTGATCTGTTGGACCAATGAT 3683
OY 3687 ATTCAAGGCATCAAGACCAGGGTGCCTGTCAGAAAGTTCTCAGACCTCTACATCTCTCAG 3746

```

|||||
Db 3684 ATTACGGCATCAAGACCAGGGTGGCCCTCAGAGTCTCCAGCCCTCCTACATCTCTCAG 3743
OY 3747 TTTATCATCATGTATAGTCTTGATGGAGAGAAGTGGCAGACTTATCGAGAAATCCACT 3806
Db 3744 TTTATCATCATGTATAGTCTTGATGGAGAGAAGTGGCAGACTTATCGAGAAATCCACT 3803
OY 3807 GGAACCTTAATGCTCTTCTTTGGCAATGTGATTCATCTGGGATAAACACATATTTT 3866
Db 3804 GGAACCTTAATGCTCTTCTTTGGCAATGTGATTCATCTGGGATAAACACATATTTT 3863
OY 3867 AACCTCCATTTATGCTCGATACATCCGTTTGCACCCAACTCATTTATAGCATTCGACG 3926
Db 3864 AACCTCCATTTATGCTCGATACATCCGTTTGCACCCAACTCATTTATAGCATTCGACG 3923
OY 3927 ACTCTTCGCATGAGTGTGATGGCTGTGATTTAAATAGTTGACAGATGCCATTGGGAATG 3986
Db 3924 ACTCTTCGCATGAGTGTGATGGCTGTGATTTAAATAGTTGACAGATGCCATTGGGAATG 3983
OY 3987 GAGAGTAAAGCAATATGATGACAGATTTACTGCTTCATCCTACTTTACCAATATGTTT 4046
Db 3984 GAGAGTAAAGCAATATGATGACAGATTTACTGCTTCATCCTACTTTACCAATATGTTT 4043
OY 4047 GCCACCTGCTCTCCTTCAAAAGCTCGACTTCACCTCCAAAGGAGAGTAATGCCCTGGAGA 4106
Db 4044 GCCACCTGCTCTCCTTCAAAAGCTCGACTTCACCTCCAAAGGAGAGTAATGCCCTGGAGA 4103
OY 4107 CCTCAGGTGAATATCCAAAGAGTGGCTGCAAGTGGACTTCCAGAGACAAATGAAGTTC 4166
Db 4104 CCTCAGGTGAATATCCAAAGAGTGGCTGCAAGTGGACTTCCAGAGACAAATGAAGTTC 4163
OY 4167 ACAGAGTAACTACTCAGGAGGAGTAATCTCTGCTTACCAGCATGTATGTGAAGAGTTC 4226
Db 4164 ACAGAGTAACTACTCAGGAGGAGTAATCTCTGCTTACCAGCATGTATGTGAAGAGTTC 4223
OY 4227 CTGATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTTTTTCAGAATGGCAAAGTA 4286
Db 4224 CTGATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTTTTTCAGAATGGCAAAGTA 4283
OY 4287 AAGGTTTTTCAGGAAATCAAGACTCTTCACACCTGTGTGAATCTCTAGACCCACCG 4346
Db 4284 AAGGTTTTTCAGGAAATCAAGACTCTTCACACCTGTGTGAATCTCTAGACCCACCG 4343
OY 4347 TTACTGACTCGTACCTTGAATTCACCCCAAGATTTGGTGCACAGATGGCCCTGAGG 4406
Db 4344 TTACTGACTCGTACCTTGAATTCACCCCAAGATTTGGTGCACAGATGGCCCTGAGG 4403
OY 4407 ATGAGGTTCTGGGCTGCGAGCAGACAGGACCTCTACTGAGGTTGGCCACTGCAGCACCTG 4466
Db 4404 ATGAGGTTCTGGGCTGCGAGCAGACAGGACCTCTACTGAGGTTGGCCACTGCAGCACCTG 4463
OY 4467 CCACTGCCGTCACTCTCCCTCCAGCTCCAGGCAAGTGTCCCTCCCTGGCTGCTTC 4526
Db 4464 CCACTGCCGTCACTCTCCCTCCAGCTCCAGGCAAGTGTCCCTCCCTGGCTGCTTC 4523
OY 4527 TACCTTTGTGTAATCCTAGCAGACACTGCTTGAAGCCTCCCTGAATTAATCAATCA 4586
Db 4524 TACCTTTGTGTAATCCTAGCAGACACTGCTTGAAGCCTCCCTGAATTAATCAATCA 4583
OY 4587 GTCCTGCATTTCTTTGGTGGGGGCCAGAGGGTGCATCCAAATTAATCAATCAATCA 4646
Db 4584 GTCCTGCATTTCTTTGGTGGGGGCCAGAGGGTGCATCCAAATTAATCAATCAATCA 4643
OY 4647 TATTTCTGCAG 4658
Db 4644 GTCGACCTGCAG 4655

```

RESULT 6
 N81544
 ID N81544 standard; DNA; 4830 BP.
 AC N81544.
 DT 04-DEC-1990 (first entry)

DE Human Factor VIII-C analog having exon 14 deleted.
 KW Human Factor VIII-C analog; exon deletion; coagulation disorders;
 KW haemophilia; ss.
 PN EP-265778-A.
 PD 04-MAY-1988.
 PF 14-OCT-1987; 115043.
 PR 15-OCT-1986; US-919153.
 PA (RORE) Rorer Int Overseas.
 PI Sarver N. Drohan W.;
 DR WPI; 88-120930/18.
 PT Human Factor VIII-C analogue free of other proteins - produced from
 PT recombinant DNA and used in the treatment of coagulation disorders.
 PS Claim 3; Page 13; 42pp; English.
 CC The genetically engineered analogue can provide a dependable and
 CC readily available therapeutic agent to be used in the treatment of
 CC haemophilia and coagulation disorders in humans or animals.
 CC See also N81543-45.
 SQ Sequence 4830 BP; 1392 A; 1091 C; 1058 G; 1289 T;

Query Match	89.5%;	Score 4322.4;	DB 1;	Length 4830;
Best Local Similarity	94.8%;	Pred. No. 0;		
Matches 4579;	Conservative 0;	Mismatches 11;	Indels 240;	Gaps 1;
OY 72 ATGCAATAGAGCTCTCCACCTGCTCTTCTGTGCGCTTTGGATTCGCTTTAGTGC 131				
Db 1 ATGCAATAGAGCTCTCCACCTGCTCTTCTGTGCGCTTTGGATTCGCTTTAGTGC 60				
OY 132 ACCAGAAGATACTACCTGGGTCAGTGGAACTGTCATGGGACTATATGCAAAAGTATCTC 191				
Db 61 ACCAGAAGATACTACCTGGGTCAGTGGAACTGTCATGGGACTATATGCAAAAGTATCTC 120				
OY 192 GGTGAGCTGCTGTGGACGCAAGATTTCCTCCTAGAGTGGCAAAATCTTTCCATTCAAC 251				
Db 121 GGTGAGCTGCTGTGGACGCAAGATTTCCTCCTAGAGTGGCAAAATCTTTCCATTCAAC 180				
OY 252 ACCTCAGTCGCTGTACAAAAGACTCTGTTGTAGAAATTCACGGATCACCTTTTCAACATC 311				
Db 181 ACCTCAGTCGCTGTACAAAAGACTCTGTTGTAGAAATTCACGGATCACCTTTTCAACATC 240				
OY 312 GCTAAGCCAGGCCACCCTGGATGGGTGTGCTAGGTCTTACCATCCAGGCTGAGGTTTAT 371				
Db 241 GCTAAGCCAGGCCACCCTGGATGGGTGTGCTAGGTCTTACCATCCAGGCTGAGGTTTAT 300				
OY 372 GATACAGTGTGCTATTACACTTAAAGACATGGCTTCCCATCCTGTCAAGTTCATGCTGTT 431				
Db 301 GATACAGTGTGCTATTACACTTAAAGACATGGCTTCCCATCCTGTCAAGTTCATGCTGTT 360				
OY 432 GGTGATCTCTACTGGAAGCTTCTGAGGAGCTGAATATGATGATCAGACCAAGTCAAAAG 491				
Db 361 GGTGATCTCTACTGGAAGCTTCTGAGGAGCTGAATATGATGATCAGACCAAGTCAAAAG 420				
OY 492 GAGAAGAAGATGATAAAGCTTCCCTGGTGGGAAGCCATACATATGTCTGGCAGGCTCTG 551				
Db 421 GAGAAGAAGATGATAAAGCTTCCCTGGTGGGAAGCCATACATATGTCTGGCAGGCTCTG 480				
OY 552 AAAGAGATGTGTCATGAGCTTGAATTCAGGCCCTCATTTGAGAGCCCTACTATATCTTCTCAT 611				
Db 481 AAAGAGATGTGTCATGAGCTTGAATTCAGGCCCTCATTTGAGAGCCCTACTATATCTTCTCAT 540				
OY 612 GTGAGCTGTGTAAGAGCTTGAATTCAGGCCCTCATTTGAGAGCCCTACTATATCTTCTCAT 671				
Db 541 GTGAGCTGTGTAAGAGCTTGAATTCAGGCCCTCATTTGAGAGCCCTACTATATCTTCTCAT 600				
OY 672 GGGAGTCTGGCCCAAGGAAAGACACAGACCTTGACAAATTTATACTATTTTGGCTGTA 731				
Db 601 GGGAGTCTGGCCCAAGGAAAGACACAGACCTTGACAAATTTATACTATTTTGGCTGTA 660				
OY 732 TTTGATGAAGGGAAGTTGGCACTGAGAAACAAAGAACTCCTTGATGACAGATAGGAT 791				
Db 661 TTTGATGAAGGGAAGTTGGCACTGAGAAACAAAGAACTCCTTGATGACAGATAGGAT 720				
OY 792 GCTGATCTGCTGGGCTGGCTTAAATGACACAGTCAATGTTATGTAAACAGGCTCT 851				

|||||
Db 721 GCTGCATCTGCTCGGGCCTGGCCTAAATGACACAGTCATGTTATGTAAACAGGTCT 780
QY 852 CTGCCAGGCTGTGATTTGGATGCGCACAGAAATCAGTCTATTTGGCATGTGATTGGAATGGGC 911
Db 781 CTGCCAGGCTGTGATTTGGATGCGCACAGAAATCAGTCTATTTGGCATGTGATTGGAATGGGC 840
~ 512 ACCACCTCTGAAGTGCACATCAATATTTCTCGAAGGTTCACACATTTCTTGTGAGGAACCAT 971
Db 841 ACCACTCTCGAAGTGCACATCAATATTTCTCGAAGGTTCACACATTTCTTGTGAGGAACCAT 900
QY 972 CGCCAGGCGTCTTGGAAATCTCGCCAATACTTCTTACTGCTCAACACTCTTGATG 1031
Db 901 CGCCAGGCGTCTTGGAAATCTCGCCAATACTTCTTACTGCTCAACACTCTTGATG 960
QY 1032 GACCTTGAAGTCTTCTACGTTTGTCTCATATCTCTCCACCAACATGATGATGGAA 1091
Db 961 GACCTTGAAGTCTTCTACGTTTGTCTCATATCTCTCCACCAACATGATGATGGAA 1020
QY 1092 GCTTATGTCAAGTAGACAGCTGTCCAGAGAAACCCCACTACGAATGAATAATGAA 1151
Db 1021 GCTTATGTCAAGTAGACAGCTGTCCAGAGAAACCCCACTACGAATGAATAATGAA 1080
QY 1152 GAAGCGAGAGACTATGATGATGATCTTACTGATTTCTGAATGATGTGTCAGTTGAT 1211
Db 1081 GAAGCGAGAGACTATGATGATGATCTTACTGATTTCTGAATGATGTGTCAGTTGAT 1140
QY 1212 GATGACAACTCTCCTCTCTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCTAAACT 1271
Db 1141 GATGACAACTCTCCTCTCTTATCCAAATTCGCTCAGTTGCCAAGAAGCATCTAAACT 1200
QY 1272 TGGGTACATTAATTTGCTGCTGAAGAGAGAGACTGGGACTATGCTCCCTTACTCTCGCC 1331
Db 1201 TGGGTACATTAATTTGCTGCTGAAGAGAGAGACTGGGACTATGCTCCCTTACTCTCGCC 1260
QY 1332 CCCGATGACAGAAGTTATAAAGTCAATATTTGAACAATGGCCCTCAGCGGATTGGTAG 1391
Db 1261 CCCGATGACAGAAGTTATAAAGTCAATATTTGAACAATGGCCCTCAGCGGATTGGTAG 1320
QY 1392 AAGTACAAAAAGTCCGATTTATGGCATACACAGATGAAACCTTTAAGACTGTGAAGCT 1451
Db 1321 AAGTACAAAAAGTCCGATTTATGGCATACACAGATGAAACCTTTAAGACTGTGAAGCT 1380
QY 1452 ATTCAGCATGAATCAGGAATTTGGGACCTTACTTTATGGGGAAGTTGGAGACACACTG 1511
Db 1381 ATTCAGCATGAATCAGGAATTTGGGACCTTACTTTATGGGGAAGTTGGAGACACACTG 1440
QY 1512 TTGATTATATTTAAGAATCAAGCAGACATATTAACATCTACCCCTCAGCGAATCACT 1571
Db 1441 TTGATTATATTTAAGAATCAAGCAGACATATTAACATCTACCCCTCAGCGAATCACT 1500
QY 1572 GATGTCGCTCTTTGTATTCAAGGAGATTACCAAAAGTGTAACCATTTGAAGGATTTT 1631
Db 1501 GATGTCGCTCTTTGTATTCAAGGAGATTACCAAAAGTGTAACCATTTGAAGGATTTT 1560
QY 1632 CCAATTTGCCAGGAGAAATATTCAAATATAAATGACAGTGAAGTGAAGTGGCCA 1691
Db 1561 CCAATTTGCCAGGAGAAATATTCAAATATAAATGACAGTGAAGTGAAGTGGCCA 1620
QY 1692 ACTAATCAGATCTCGGTGCTGACCCGCTATTAATTAATGAGAGAGA 1751
Db 1621 ACTAATCAGATCTCGGTGCTGACCCGCTATTAATTAATGAGAGAGA 1680
QY 1752 GATCTAGCTTCAGGACTCATTTGCCCCCTCTCTCATCTGCTACAAAGAATCTGATCAA 1811
Db 1681 GATCTAGCTTCAGGACTCATTTGCCCCCTCTCTCATCTGCTACAAAGAATCTGATCAA 1740
QY 1812 AGAGAAACAGATATGTCAAGAGAGAAATGTCATCTGTTTCTGATTTGATGAG 1871
Db 1741 AGAGAAACAGATATGTCAAGAGAGAAATGTCATCTGTTTCTGATTTGATGAG 1800
QY 1872 AACCGAAGCTGCTACCTCACAGAGATATACACGCTTCTCCCAATCCAGCTGAGTG 1931
|||||

Db 1801 AACCGAAGCTGCTACCTCACAGAGAAATATACAAAGCTTTCTCCCAATCCAGCTGAGTG 1860
QY 1932 CAGCTTGAGATCCAGAGTTCCAAAGCCTCCAAACATCATGACAGCATCAATGCTATGTT 1991
Db 1861 CAGCTTGAGATCCAGAGTTCCAAAGCCTCCAAACATCATGACAGCATCAATGCTATGTT 1920
QY 1992 TTGATAGTTTGCAGTTGTCAGTTTGTGTCATGAGTGGCATACTGTTACATTTCTAAGC 2051
Db 1921 TTGATAGTTTGCAGTTGTCAGTTTGTGTCATGAGTGGCATACTGTTACATTTCTAAGC 1980
QY 2052 ATTGAGCAGAGACTGACTTCCCTTCTGTCCTTCTCTCTGATATACCTTCAACACAAA 2111
Db 1981 ATTGAGCAGAGACTGACTTCCCTTCTGTCCTTCTCTCTGATATATCTTCAACACAAA 2040
QY 2112 ATGCTATGAAGACACACATCAACCTATTCCTATTCCTCAGAGAAACTGTTCTATGTCG 2171
Db 2041 ATGCTATGAAGACACACATCAACCTATTCCTATTCCTCAGAGAAACTGTTCTATGTCG 2100
QY 2172 ATGAAAACCCAGGCTATGATTTCTGGGTCGCACAACTCAGACTTTCGGAACAGAGGC 2231
Db 2101 ATGAAAACCCAGGCTATGATTTCTGGGGTGCACAACTCAGACTTTCGGAACAGAGGC 2160
QY 2232 ATGACCGCTTACTGAAGTTTCTAGTTGTGACAAGAACTGCTGATTTATACAGAGAC 2291
Db 2161 ATGACCGCTTACTGAAGTTTCTAGTTGTGACAAGAACTGCTGATTTATACAGAGAC 2220
QY 2292 AGTTATGAAGATTTTCAGCATTAATTTGCTGAGTAAACAAATGCAATTTGAACCAAGAGC 2351
Db 2221 AGTTATGAAGATTTTCAGCATTAATTTGCTGAGTAAACAAATGCAATTTGAACCAAGAGC 2280
QY 2352 TTCTCCAG----- 2361
Db 2281 TTCTCCAGAAATTCACTATTTGATCCTCTTCTGGAATAACCACTATGTTACTCAGATA 2340
QY 2361 ----- 2361
Db 2341 CCAAAAGAGAGTGGAATCCCAAGAGAGTCAACGAAAAAAGAGCTTTTAAGAAAAAG 2400
QY 2361 ----- 2361
Db 2401 GATACCATTTTGTCCCTGAAGCCTTGTGAAGCAATCATGCAATAGCAGCAATAATGAG 2460
QY 2361 ----- 2361
Db 2461 GGACAAATAGACCCGAATATGAAGTCACTCGGGCACAGCAAGGTAGACTGAAGGCTG 2520
QY 2361 -----AACCACACAGTCTTGAACCGCATCAACGGGAATACTCGTACTACTT 2411
Db 2521 TGCTCTCAAAACCCAGTCTTGAACCGCATCAACGGGAATACTCGTACTACTT 2580
QY 2412 CAGTCAGATCAAGAGAAATGACTATGATGATTAACCATATCAGTTGAATGAAGAGAA 2471
Db 2581 CAGTCAGATCAAGAGAAATGACTATGATGATTAACCATATCAGTTGAATGAAGAGAA 2640
QY 2472 GATTTGACATTTATGATGAGATGAAAAATCAGAGCCCCCGCAGCTTTCAAAAAGAAACA 2531
Db 2641 GATTTGACATTTATGATGAGATGAAAAATCAGAGCCCCCGCAGCTTTCAAAAAGAAACA 2700
QY 2532 CGACACTATTTTATTTGCTGCACTGAGAGGCTCTGGGATTAATGGGATGAGTAGCTCCCA 2591
Db 2701 CGACACTATTTTATTTGCTGCACTGAGAGGCTCTGGGATTAATGGGATGAGTAGCTCCCA 2760
QY 2592 CATGTTCTAAGAAACAGGGCTCAGAGTGGAGTGTCCCTCAGTTCAGAAGAAAGTTGTTTC 2651
Db 2761 CATGTTCTAAGAAACAGGGCTCAGAGTGGAGTGTCCCTCAGTTCAGAAGAAAGTTGTTTC 2820
QY 2652 CAGGAATTTACTGATGGCTCTTACTCAGCGCTTATACCGTGGAGAACTAATGAACAT 2711
Db 2821 CAGGAATTTACTGATGGCTCTTACTCAGCGCTTATACCGTGGAGAACTAATGAACAT 2880
QY 2712 TTGGGACTCTGGGGCCATATATTAAGAGCAGAGTGAAGATATATCATGTTAATCTTTC 2771
Db 2881 TTGGGACTCTGGGGCCATATATTAAGAGCAGAGTGAAGATATATCATGTTAATCTTTC 2940

QY 2772 AGAATCAGGCGCTCGTCCCTATTCCTCTATTCAGCCTATTTCTTATGAGAGAT 2831
 |||||
 Db 2941 AGAAATCAGGCGCTCGTCCCTATTCCTCTATTCAGCCTATTTCTTATGAGAGAT 3000
 |||||
 QY 2932 CAGAGCAAGGAGCAGAACCTAGAAAAAACTTTGTCAAGCCTAATGAACCAAACTTAC 2891
 |||||
 Db 3001 CAGAGCAAGGAGCAGAACCTAGAAAAAACTTTGTCAAGCCTAATGAACCAAACTTAC 3060
 |||||
 QY 2892 TTTTGGAAAGTGCAACATCATATGGACCCCACTAAGATGAGTTTGTACTGCAAGCCTGG 2951
 |||||
 Db 3061 TTTTGGAAAGTGCAACATCATATGGACCCCACTAAGATGAGTTTGTACTGCAAGCCTGG 3120
 |||||
 QY 2952 GCTTATTTCTCTGATGTGACCTGGAAAAAGATGTGCACTCAGGCGCTGATTGGACCCCTT 3011
 |||||
 Db 3121 GCTTATTTCTCTGATGTGACCTGGAAAAAGATGTGCACTCAGGCGCTGATTGGACCCCTT 3180
 |||||
 QY 3012 CTGGTCTGCCACACTAACACACTGAAACCCTGCTCATGGGAGACAGACAGTACAGGAA 3071
 |||||
 Db 3181 CTGGTCTGCCACACTAACACACTGAAACCCTGCTCATGGGAGACAGACAGTACAGGAA 3240
 |||||
 QY 3072 TTGCTCTGTTTTCACCATCTTTGATGAGACCAAAAGCTGTACTTCACTGAATAATATG 3131
 |||||
 Db 3241 TTGCTCTGTTTTCACCATCTTTGATGAGACCAAAAGCTGTACTTCACTGAATAATATG 3300
 |||||
 QY 3132 GAAAGAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAT 3191
 |||||
 Db 3301 GAAAGAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAT 3360
 |||||
 QY 3192 TATGCTTCCATGCAATCAATGGCTACATATATGATGATACACTACCTGGCTTAGTAAATGGCT 3251
 |||||
 Db 3361 TATGCTTCCATGCAATCAATGGCTACATATATGATGATACACTACCTGGCTTAGTAAATGGCT 3420
 |||||
 QY 3252 CAGGATCAAGAGATTCGATGATCTGCTCAGCATGGGAGCAATGAAGACATCCATTTCT 3311
 |||||
 Db 3421 CAGGATCAAGAGATTCGATGATCTGCTCAGCATGGGAGCAATGAAGACATCCATTTCT 3480
 |||||
 QY 3312 ATTCATTTCACTGACATGTGTCACTGTACGAAAAAAAGAGAGTATTAATGGCACTG 3371
 |||||
 Db 3481 ATTCATTTCACTGACATGTGTCACTGTACGAAAAAAAGAGAGTATTAATGGCACTG 3540
 |||||
 QY 3372 TACAATCTCTATCCAGTGTGTTTGGAGACAGTGAAGATGTTACCATCCAAAGCTGGAATT 3431
 |||||
 Db 3541 TACAATCTCTATCCAGTGTGTTTGGAGACAGTGAAGATGTTACCATCCAAAGCTGGAATT 3600
 |||||
 QY 3432 TGGCGGGTGAATGCCCTTATTTGGCGAGCATCTACATGCTGGATGAGCACACTTTTCTG 3491
 |||||
 Db 3601 TGGCGGGTGAATGCCCTTATTTGGCGAGCATCTACATGCTGGATGAGCACACTTTTCTG 3660
 |||||
 QY 3492 GTGTACAGCAATAGTGTGAGACTCCCTGGGAATGGCTTGTGACACATTAGAGATTTT 3551
 |||||
 Db 3661 GTGTACAGCAATAGTGTGAGACTCCCTGGGAATGGCTTGTGACACATTAGAGATTTT 3720
 |||||
 QY 3552 CAGATTACAGCTTCAGGACAATATGACAGTGGGGCCCCAAAGCTGGCCAGACTTCATTAT 3611
 |||||
 Db 3721 CAGATTACAGCTTCAGGACAATATGACAGTGGGGCCCCAAAGCTGGCCAGACTTCATTAT 3780
 |||||
 QY 3612 TCCGGATCAATCAATGCCCTGGAGACCAAGAGCCCTTTTCTTGATCAAGGTGATCTG 3671
 |||||
 Db 3781 TCCGGATCAATCAATGCCCTGGAGACCAAGAGCCCTTTTCTTGATCAAGGTGATCTG 3840
 |||||
 QY 3672 TTGGACCAATGATTATTCAGGCGCATCAAGACCCAGGGTCCCGCTCAGAGTTCTCCAGC 3731
 |||||
 Db 3841 TTGGACCAATGATTATTCAGGCGCATCAAGACCCAGGGTCCCGCTCAGAGTTCTCCAGC 3900
 |||||
 QY 3732 CTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGAAGTGGCAGACTTAT 3791
 |||||
 Db 3901 CTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGAAGTGGCAGACTTAT 3960
 |||||
 QY 3792 CGAGGAAATTCACACTGGAACCTTAATGCTTCTTTGGCAATGTGATTTACTTGGGATA 3851
 |||||
 Db 3961 CGAGGAAATTCACACTGGAACCTTAATGCTTCTTTGGCAATGTGATTTACTTGGGATA 4020
 |||||

QY 3852 AAACACAATATTTTAAACCCTCCAAATATATGCTCGATACATCCGTTTGCACCCAACTCAT 3911
 |||||
 Db 4021 AAACACAATATTTTAAACCCTCCAAATATATGCTCGATACATCCGTTTGCACCCAACTCAT 4080
 |||||
 QY 3912 TATAGCATTCGAGCAGCAGCTCTTCGCATGAGATTTGATGGCGTGTGATTTAAATAGTTGCAGC 3971
 |||||
 Db 4081 TATAGCATTCGAGCAGCAGCTCTTCGCATGAGATTTGATGGCGTGTGATTTAAATAGTTGCAGC 4140
 |||||
 QY 3972 ATGCCATTGGGAATGGAGAGTAAAGCAATATCAGATGACAGATTACTGCTTCATCTTAC 4031
 |||||
 Db 4141 ATGCCATTGGGAATGGAGAGTAAAGCAATATCAGATGACAGATTACTGCTTCATCTTAC 4200
 |||||
 QY 4032 TTTACCAATATGTTTGGCACCTGGTCTCTCTTCAAAAGCTGCACTTCACTCCAAAGGAGG 4091
 |||||
 Db 4201 TTTACCAATATGTTTGGCACCTGGTCTCTCTTCAAAAGCTGCACTTCACTCCAAAGGAGG 4260
 |||||
 QY 4092 AGTAATGCTGGAGACCTCAGGTGATATAATCCAAAGAGTGGCTGCAAGTGGACTTCCAG 4151
 |||||
 Db 4261 AGTAATGCTGGAGACCTCAGGTGATATAATCCAAAGAGTGGCTGCAAGTGGACTTCCAG 4320
 |||||
 QY 4152 AAGACAATGAAGTCAACAGAGTACTACTCAGGAGTAAATCTCTGCTTACAGCATG 4211
 |||||
 Db 4321 AAGACAATGAAGTCAACAGAGTACTACTCAGGAGTAAATCTCTGCTTACAGCATG 4380
 |||||
 QY 4212 TATGTGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTCTTTT 4271
 |||||
 Db 4381 TATGTGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTCTTTT 4440
 |||||
 QY 4272 CAGATGGCAAGTAAAGGTTTTTCAGGGAATCAAGACTCTTCAACACTGTGTGAC 4331
 |||||
 Db 4441 CAGATGGCAAGTAAAGGTTTTTCAGGGAATCAAGACTCTTCAACACTGTGTGAC 4500
 |||||
 QY 4332 TCTCTAGACCCACCGTTACTGACTCGCTACCTTGAATTCACCCCCAGAGTTGGGTGAC 4391
 |||||
 Db 4501 TCTCTAGACCCACCGTTACTGACTCGCTACCTTGAATTCACCCCCAGAGTTGGGTGAC 4560
 |||||
 QY 4392 CAGATTGCCCTGAGGATGAGGTTCCTGGGCTGCGAGGACAGGACCTCTACTGAGGTTGG 4451
 |||||
 Db 4561 CAGATTGCCCTGAGGATGAGGTTCCTGGGCTGCGAGGACAGGACCTCTACTGAGGTTGG 4620
 |||||
 QY 4452 CCACCTGACGACCTGCGACCTGCCGTACCTCTCCCTCTCAGCTCCAGGGCAGTGTCCCT 4511
 |||||
 Db 4621 CCACCTGACGACCTGCGACCTGCCGTACCTCTCCCTCTCAGCTCCAGGGCAGTGTCCCT 4680
 |||||
 QY 4512 CCCTGGCTTGGCTTCTACCTTTGTGCTAAATCTTAGACAGACACTGCTTGAAGCTTCTG 4571
 |||||
 Db 4681 CCCTGGCTTGGCTTCTACCTTTGTGCTAAATCTTAGACAGACACTGCTTGAAGCTTCTG 4740
 |||||
 QY 4572 AATTAATATCATCAGTCCCTGCATTTCTTGTGGGGGGCCAGAGGTTGCATCCAAATT 4631
 |||||
 Db 4741 AATTAATATCATCAGTCCCTGCATTTCTTGTGGGGGGCCAGAGGTTGCATCCAAATT 4800
 |||||
 QY 4632 AACTTAATCTTACCTATTTTCTGCAGCTG 4661
 |||||
 Db 4801 AACTTAATCTTACCTATTTTCTGCAGCTG 4830
 |||||

RESULT 7
 N90654
 ID N90654 standard; DNA; 4275 BP.
 AC N90654;
 DT 26-JUN-1990 (first entry)
 DE DNA encoding 740 Arg-1649 Glu human Factor VIII:C
 KW Human Factor VIII:C; Ad.RE.neo; 740 Arg-1649 Glu Factor VIII:C;
 OS Homo sapiens.
 FH Key
 FT cds
 FT 1..4275
 Location/Qualifiers
 /*tag= a
 EP-306968-A.
 PD 15-MAR-1989.
 PF 09-SEP-1988; 114769.
 PR 08-APR-1988; JP-085454.

PA (KAGA) Chemo-Sero-Therap (Telj).
PI Sugiyama T, Masuda K, Tajima Y, Yonemura H;
DR WPI: 89-078467/11.
DR P-PSDB: P91165.
PT Prodn. of recombinant human Factor-VIII-C -
PT using animal cells transformed with a vector contg. the gene for
PS Factor VIII:C and a promoter
PS Figure 1(1) - 1(13); 32pp; English.
CC When translated, Arg-740 of the carboxyl terminus of the H chain is
CC directly bonded by a peptide bond to Glu-1649 of the amino terminus of
CC L chain. It is used to transform animal cells so that they produce
CC human Factor VIII:C. A pref. expression vector is plasmid Ad.RE.neo.
CC The expression vector has at least one promoter upstream of n90654.
CC The transformants can constantly and continuously produce human Factor
CC VIII:C in high yield on a commercial scale. The human Factor VIII:C so
CC produced is considered to corresp. to the smallest species of active and
CC intact Factor VIII:C molecules in the human blood plasma. It is useful
CC for treating haemophilia A patients.
SQ Sequence 4275 BP; 1245 A; 941 C; 945 G; 1144 T;

Query Match 87.3%; Score 4216.6; DB 1; Length 4275;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 4271; Conservative 0; Mismatches 4; Indels 42; Gaps 1;

QY 129 GCCACCAGAGATCTACCTGGGTCAGTGAAGTGCATGGGACTATATGCAAGTGAT 188
DB 1 GCCACCAGAGATCTACCTGGGTCAGTGAAGTGCATGGGACTATATGCAAGTGAT 60
QY 189 CTCGGTGAGCTGCTGTGACGCAAGATTCTCTCTAGAGTGCACAAATCTTTCCATT 248
DB 61 CTCGGTGAGCTGCTGTGACGCAAGATTCTCTCTAGAGTGCACAAATCTTTCCATT 120
QY 249 AACACCTCAGTGTGTACAAAAGACTGTGTGTAGAAATTCACGGATCACCTTTTCAAC 308
DB 121 AACACCTCAGTGTGTACAAAAGACTGTGTGTAGAAATTCACGGATCACCTTTTCAAC 180
QY 309 ATCGCTAAGCCAAAGGCCACCTGTGATGGGTCTGTAGGTCTCTACCATCCAGGCTGAGGTT 368
DB 181 ATCGCTAAGCCAAAGGCCACCTGTGATGGGTCTGTAGGTCTCTACCATCCAGGCTGAGGTT 240
QY 369 TATGATACAGTGTCTATTACACTTAAGAATGCTTCCATCTCTGACCTTCATGCT 428
DB 241 TATGATACAGTGTCTATTACACTTAAGAATGCTTCCATCTCTGACCTTCATGCT 300
QY 429 GTTGGTGTATCTTACTGGAAGCTTCTGAGGAGCTGAATATGATCAGACCACTCAA 488
DB 301 GTTGGTGTATCTTACTGGAAGCTTCTGAGGAGCTGAATATGATCAGACCACTCAA 360
QY 489 AGGGAGAAAGAGATGATAAGTCTTCCCTGTGGAAGCCATACATATGTCTGGCAGGTC 548
DB 361 AGGGAGAAAGAGATGATAAGTCTTCCCTGTGGAAGCCATACATATGTCTGGCAGGTC 420
QY 549 CTGAAGAGAGATGCTCAATGGCTCTGACCCACTGTGCTTACTACTATCTTCT 608
DB 421 CTGAAGAGAGATGCTCAATGGCTCTGACCCACTGTGCTTACTACTATCTTCT 480
QY 609 CATGTGACCTGTGTAAGACTTGAATTGAGGCTCATTTGAGCCCTACTAGATGTAGA 668
DB 481 CATGTGACCTGTGTAAGACTTGAATTGAGGCTCATTTGAGCCCTACTAGATGTAGA 540
QY 669 GAAGGAGTCTGGCCAAAGAAAGACACAGACCTTGACAAATTTACTACTTTTGTCT 728
DB 541 GAAGGAGTCTGGCCAAAGAAAGACACACACCTTGACAAATTTACTACTTTTGTCT 600
QY 729 GTATTGTGAAGGAAAGTGTGCACTCAGAAACAAAGACTCTGTATGACAGATAGG 788
DB 601 GTATTGTGAAGGAAAGTGTGCACTCAGAAACAAAGACTCTGTATGACAGATAGG 660
QY 789 GATGCTGCATCTGCTCGGCTGAGCTTAAATGACACAGTCAATGTTATGTAACAGG 848
DB 661 GATGCTGCATCTGCTCGGCTGAGCTTAAATGACACAGTCAATGTTATGTAACAGG 720

QY 849 TCTCTGCCAGGCTCTGATTGATGACACAGGAATCAGTCTATTGGCATGTGATTGGAATG 908
DB 721 TCTCTACCAAGTCTGATTGATGACACAGGAATCAGTCTATTGGCATGTGATTGGAATG 780
QY 909 GGCACCACTCTGAAAGTGCACATCATATTTCTGGAAGTGCACACATTTCTTGAGGAAC 968
DB 781 GGCACCACTCTGAAAGTGCACATCATATTTCTGGAAGTGCACACATTTCTTGAGGAAC 840
QY 969 CATGCCAGGCTCTTGAAATCTGCCAATTAATTTCTTACTGCTCAAACTCTTG 1028
DB 841 CATGCCAGGCTCTTGAAATCTGCCAATTAATTTCTTACTGCTCAAACTCTTG 900
QY 1029 ATGACCTTGACAGTTTCTACTGTTTGTCAATCTCTTCCACCAACATGATGCGATG 1088
DB 901 ATGACCTTGACAGTTTCTACTGTTTGTCAATCTCTTCCACCAACATGATGCGATG 960
QY 1089 GAAGCTTATGCAAAAGTAGACAGCTGTCCAGAGGAACCCCACTACGAATGAATAAT 1148
DB 961 GAAGCTTATGCAAAAGTAGACAGCTGTCCAGAGGAACCCCACTACGAATGAATAAT 1020
QY 1149 GAAGAAGCGAAGACTATGATGATGATCTTACTGATTTCTGAATGATGTGTCAGGTTT 1208
DB 1021 GAAGAAGCGAAGACTATGATGATGATCTTACTGATTTCTGAATGATGTGTCAGGTTT 1080
QY 1209 GATGATGACAACTCTCTCTTATCCAAATCGCTCAGTGTCCAAAGAACCTTAA 1268
DB 1081 GATGATGACAACTCTCTCTTATCCAAATCGCTCAGTGTCCAAAGAACCTTAA 1140
QY 1269 ACTGGGTACATTAATGCTGTGGAAGAGAGAGAGTGGACTATGCTCTAGTCTC 1328
DB 1141 ACTGGGTACATTAATGCTGTGGAAGAGAGAGAGTGGACTATGCTCTAGTCTC 1200
QY 1329 GCCCCGATGACAGAAAGTATTAAGTCAATATTGAAACAATGCCCTCAGCGGATGTT 1388
DB 1201 GCCCCGATGACAGAAAGTATTAAGTCAATATTGAAACAATGCCCTCAGCGGATGTT 1260
QY 1389 AGAAGTACAAAAGTCCGATTTATGCAATACAGATGAACCTTTAAGACTCTGTA 1448
DB 1261 AGAAGTACAAAAGTCCGATTTATGCAATACAGATGAACCTTTAAGACTCTGTA 1320
QY 1449 GCTATTACAGATGATCAGAAATCTTGGACCTTACTTTATGGGGAAGTGGACACACA 1508
DB 1321 GCTATTACAGATGATCAGAAATCTTGGACCTTACTTTATGGGGAAGTGGACACACA 1380
QY 1509 CTGTGATTTATTTAAGATCAAGCAAGCACCATATTAACATCTACCCCTCAGGAATC 1568
DB 1381 CTGTGATTTATTTAAGATCAAGCAAGCACCATATTAACATCTACCCCTCAGGAATC 1440
QY 1569 ACTGATGTCGCTCTTGTATTCAGAGATTAACCAAAAGGTGTAACATTTGAAGAT 1628
DB 1441 ACTGATGTCGCTCTTGTATTCAGAGATTAACCAAAAGGTGTAACATTTGAAGAT 1500
QY 1629 TTTCCAATTTCTGCCAGAGAAATATCAAAATTAATTAATGACAGTGAATGGG 1688
DB 1501 TTTCCAATTTCTGCCAGAGAAATATCAAAATTAATTAATGACAGTGAATGGG 1560
QY 1689 CCAACTAATCAGATCTCTGGTGCCTGACCCGCTATTACTCTAGTTTGTATATGAG 1748
DB 1561 CCAACTAATCAGATCTCTGGTGCCTGACCCGCTATTACTCTAGTTTGTATATGAG 1620
QY 1749 AGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGTCTACAAAGAACTGTAGAT 1808
DB 1621 AGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCTCATCTGTCTACAAAGAACTGTAGAT 1680
QY 1809 CAAAGAGAAACAGATTAATGTCAGACAAAGAGAAATGTCATCTGTTTCTGTATTTGAT 1868
DB 1681 CAAAGAGAAACAGATTAATGTCAGACAAAGAGAAATGTCATCTGTTTCTGTATTTGAT 1740
QY 1869 GAGAACCGAAGCTGTACTCTACAGAGAAATATACAAAGCTTTCTCCCAATCAGCTGGA 1928
DB 1741 GAGAACCGAAGCTGTACTCTACAGAGAAATATACAAAGCTTTCTCCCAATCAGCTGGA 1800
QY 1929 GTGCAGCTTGAGGATCCAGAGTTCCAAAGCTTCCAAACATCATGACAGATCAATGGCTAT 1988

Db 1801 GTGAGCTTGAGGATCCAGAGTCCAGCCCTCAACATCATGCACAGCATCAATGGCTAT 1860
QY 1989 GTTTTGTAGTTTGCAAGTTCAGTTGTTGTTGTCATGAGGTGGCATACTGGTACATTCTA 2048
Db 1861 GTTTTGTAGTTTGCAAGTTCAGTTGTTGTTGTCATGAGGTGGCATACTGGTACATTCTA 1920
QY 2049 AGCATTGAGCAGACAGACTGACTTCCCTTCTGTCCTTCTCTGATATACCTTCAACAC 2108
Db 1921 AGCATTGAGCAGACAGACTGACTTCCCTTCTGTCCTTCTCTGATATACCTTCAACAC 1980
QY 2109 AAAATGCTATGAAGACACACACTCACCCTATTCOCATTCAGGAGAAACTGCTTCATG 2168
Db 1981 AAAATGCTATGAAGACACACTCACCCTATTCOCATTCAGGAGAAACTGCTTCATG 2040
QY 2169 TCGATGAAACCAGGTCTATGGATTCTGGGGTCCACAACTCAGACTTTCGGAACAGA 2228
Db 2041 TCGATGAAACCAGGTCTATGGATTCTGGGGTCCACAACTCAGACTTTCGGAACAGA 2100
QY 2229 GGCATGACCGCCTTACTGAAGTTTCTAGTTGTGACAGAACACTGGTATTTACGAG 2288
Db 2101 GGCATGACCGCCTTACTGAAGTTTCTAGTTGTGACAGAACACTGGTATTTACGAG 2160
QY 2289 GACAGTATGAAGATATTTTCAGCATACTGCTGAGTAAACAATGCCATTGAACCAAGA 2348
Db 2161 GACAGTATGAAGATATTTTCAGCATACTGCTGAGTAAACAATGCCATTGAACCAAGA 2220
QY 2349 AGCTTCCCAAGAACCCAGTCTTGAACGCCATCAAGGGAATTACTGCTACTACT 2408
Db 2221 -----GAAATACTCGTACTACT 2238
QY 2409 CTTCAGTCAGATCAAGAGAAATTGACATATGATGATACCATATCAGTTGAAATGAAGAAG 2468
Db 2239 CTTCAGTCAGATCAAGAGAAATTGACATATGATGATACCATATCAGTTGAAATGAAGAAG 2298
QY 2469 GAAGATTTGACATTTATGATGAGAGTGAATAATCAGAGCCCCCGCAGCTTTCAAAAGAAA 2528
Db 2299 GAAGATTTGACATTTATGATGAGAGTGAATAATCAGAGCCCCCGCAGCTTTCAAAAGAAA 2358
QY 2529 ACACGACACTATTTTATGCTGACAGTGAAGGCTCTGGGATTATGGATGAGTAGCTCC 2588
Db 2359 ACACGACACTATTTTATGCTGACAGTGAAGGCTCTGGGATTATGGATGAGTAGCTCC 2418
QY 2589 CCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTTCCTCAGTTCAAGAAAGTTGTT 2648
Db 2419 CCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTTCCTCAGTTCAAGAAAGTTGTT 2478
QY 2649 TTCCAGAAATTTACTGATGGCTCCCTTACTCAGCCCTTATACCGTGAAGAACTAATGAA 2708
Db 2479 TTCCAGAAATTTACTGATGGCTCCCTTACTCAGCCCTTATACCGTGAAGAACTAATGAA 2538
QY 2709 CATTTGGGACTCCTGGGCCATATATAAGACAGAGTTGAAGATAATATCATGTTAACT 2768
Db 2539 CATTTGGGACTCCTGGGCCATATATAAGACAGAGTTGAAGATAATATCATGTTAACT 2598
QY 2769 TTCCAGAAATCAGGCTCTGCTCCCTATTCCTTCTATCTAGCCTTATTTCTTATGAGAA 2828
Db 2599 TTCCAGAAATCAGGCTCTGCTCCCTATTCCTTCTATCTAGCCTTATTTCTTATGAGAA 2658
QY 2829 GATCAGAGGCAAGAGCAGAACTTGAATAAATCTTGTCAAGCCTAATGAACCAAAACT 2888
Db 2659 GATCAGAGGCAAGAGCAGAACTTGAATAAATCTTGTCAAGCCTAATGAACCAAAACT 2718
QY 2889 TACTTTTGAAGTGAACATCATATGCAACCCACTAAAGATGAGTTTGACTGCAAGCC 2948
Db 2719 TACTTTTGAAGTGAACATCATATGCAACCCACTAAAGATGAGTTTGACTGCAAGCC 2778
QY 2949 TGGGCTTATTTCTGATGTTGACCTGGAATAAGATGTGCACTCAGCCCTGATTTGAGCC 3008
Db 2779 TGGGCTTATTTCTGATGTTGACCTGGAATAAGATGTGCACTCAGCCCTGATTTGAGCC 2838
QY 3009 CTTCGTCTGACACACTAAGCACTGAACCCCTCATGGAGACAAGTAGACTACAG 3068
Db 3009 CTTCGTCTGACACACTAAGCACTGAACCCCTCATGGAGACAAGTAGACTACAG 3068

Db 2839 CTTCGTCTGACACACTAAGCACTGAACCCCTGCTCATGGAGACAAGTAGACTACAG 2898
QY 3069 GAATTGCTCTGTTTTCACCATCTTTTGTATGAGAGCAAAAGCTGGTACTTCACTGAAAT 3128
Db 2899 GAATTGCTCTGTTTTCACCATCTTTTGTATGAGAGCAAAAGCTGGTACTTCACTGAAAT 2958
QY 3129 ATGGAAGAAACTGAGGGCTCCCTGCAATATCCAGATGGAAGATCCACTTTTAAAGAG 3188
Db 2959 ATGGAAGAAACTGAGGGCTCCCTGCAATATCCAGATGGAAGATCCACTTTTAAAGAG 3018
QY 3189 AATTATCCCTTCCATGCAATCAATGGCTACATATGATACACTACCTGGCTTAGTAATG 3248
Db 3019 AATTATCCCTTCCATGCAATCAATGGCTACATATGATACACTACCTGGCTTAGTAATG 3078
QY 3249 GCTCAGATCAAGAGATTGATGGTATCTGCTCAGCATGGGCGCAATGAAAACATCCAT 3308
Db 3079 GCTCAGATCAAGAGATTGATGGTATCTGCTCAGCATGGGCGCAATGAAAACATCCAT 3138
QY 3309 TCTATTCATTTACGTGACATGCTGTTCACTGTACGAAAAAAGAGAGTATATAATGCA 3368
Db 3139 TCTATTCATTTACGTGACATGCTGTTCACTGTACGAAAAAAGAGAGTATATAATGCA 3198
QY 3369 CTGTACAATCTCTATCCAGGTGTTTGTAGACAGTGGAAATGTTACCATCCAAGCTGGA 3428
Db 3199 CTGTACAATCTCTATCCAGGTGTTTGTAGACAGTGGAAATGTTACCATCCAAGCTGGA 3258
QY 3429 ATTTGGCGGGTGAATGCCCTATTTGGCGAGCATCTACATGCTGGGATGAGCACACTTTT 3488
Db 3259 ATTTGGCGGGTGAATGCCCTATTTGGCGAGCATCTACATGCTGGGATGAGCACACTTTT 3318
QY 3489 CTGCTGTACAGCAATTAAGTGTAGACTCCCTGGGAATGGCTTCTGACACATTAGAGAT 3548
Db 3319 CTGCTGTACAGCAATTAAGTGTAGACTCCCTGGGAATGGCTTCTGACACATTAGAGAT 3378
QY 3549 TTTCAGATTACAGCTTCAGAGCAATATGAGAGTGGGCCCAAGCTGGCGAGACTTCAT 3608
Db 3379 TTTCAGATTACAGCTTCAGAGCAATATGAGAGTGGGCCCAAGCTGGCGAGACTTCAT 3438
QY 3609 TATTCGGATCAATCAATGCCCTGGAGACCAAGAGCCCTTTCTTGGATCAAGGTGAT 3668
Db 3439 TATTCGGATCAATCAATGCCCTGGAGACCAAGAGCCCTTTCTTGGATCAAGGTGAT 3498
QY 3669 CTGTTGGACCAATGATTTATTCACGGGATCAAGACCCAGGGTCCCGTCAGAAAGTTCTCC 3728
Db 3499 CTGTTGGACCAATGATTTATTCACGGGATCAAGACCCAGGGTCCCGTCAGAAAGTTCTCC 3558
QY 3729 AGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGGTGGCAGACT 3788
Db 3559 AGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGGTGGCAGACT 3618
QY 3789 TATCGAGAAATTCACCTGGAACCTTAATGCTCTTCTTGGCAATGTGGAATTCATCTGG 3848
Db 3619 TATCGAGAAATTCACCTGGAACCTTAATGCTCTTCTTGGCAATGTGGAATTCATCTGG 3678
QY 3849 ATAAACACAAATATTTTAAACCTCCAAATTAATGCTGCTGCAATCCGTTTGCACCCAACT 3908
Db 3679 ATAAACACAAATATTTTAAACCTCCAAATTAATGCTGCTGCAATCCGTTTGCACCCAACT 3738
QY 3909 CATATAGCAATTCGACAGACTCTTCGCAATGAGTGAATGGGCTGTGAATTAATAGTTGC 3968
Db 3739 CATATAGCAATTCGACAGACTCTTCGCAATGAGTGAATGGGCTGTGAATTAATAGTTGC 3798
QY 3969 AGCATGCCAATGGGAATGAGAGTAAAGCAATATCAGATGCACAGATTAATGCTTCAATCC 4028
Db 3799 AGCATGCCAATGGGAATGAGAGTAAAGCAATATCAGATGCACAGATTAATGCTTCAATCC 3858
QY 4029 TACTTTACCAATATGTTTGGCAACCTGGTCTCTTCAAAAGCTGCACCTCAAGGG 4088
Db 3859 TACTTTACCAATATGTTTGGCAACCTGGTCTCTTCAAAAGCTGCACCTCAAGGG 3918
QY 4089 AGGAGTAATGCTTGAGACCTCAGGTGAATATCCAAAAGAGTGGCTGCAAGTGACTTC 4148
Db 3919 AGGAGTAATGCTTGAGACCTCAGGTGAATATCCAAAAGAGTGGCTGCAAGTGACTTC 3978

QY 1449 GCTATTGACATGAATCAGAAATCTTGGACCTTACTTTATGGGGAAGTTGGAGACACA 1508
|||||
Db 1324 GCTATTGACATGAATCAGAAATCTTGGACCTTACTTTATGGGGAAGTTGGAGACACA 1383
QY 1509 CTGTTGATTATATTTAAGAATCAAGCAAGCAGACCATATATACTATACCCTCAGCGAATC 1568
|||||
Db 1384 CTGTTGATTATATTTAAGAATCAAGCAAGCAGACCATATATACTATACCCTCAGCGAATC 1443
QY 1569 ACTGATGTCCTGCTTTGTTGATTCAGAGAGATTAACCAAAAGTGTAACATTTGAAGGAT 1628
|||||
Db 1444 ACTGATGTCCTGCTTTGTTGATTCAGAGAGATTAACCAAAAGTGTAACATTTGAAGGAT 1503
QY 1629 TTTCGAATTTGCGCAGAGAAATATTCAAAATATAATGACAGTGAAGTGAAGATGGG 1688
|||||
Db 1504 TTTCGAATTTGCGCAGAGAAATATTCAAAATATAATGACAGTGAAGTGAAGATGGG 1563
QY 1689 CCAACTAAATCAGATCCTCGGTGCGCTGACCCGCTATTACTCTAGTTTTCGTTAATATGAG 1748
|||||
Db 1564 CCAACTAAATCAGATCCTCGGTGCGCTGACCCGCTATTACTCTAGTTTTCGTTAATATGAG 1623
QY 1749 AGAGATCTAGCTTCAGGACATGTTGGCCCTCTCTCATCTGCTACAAAGAAATCTGTAGAT 1808
|||||
Db 1624 AGAGATCTAGCTTCAGGACATGTTGGCCCTCTCTCATCTGCTACAAAGAAATCTGTAGAT 1683
QY 1809 CAAAGAGAAACCAGATATGTTCAGACAAGAGAAATGTCAATCCTGTTTTCTGTATTGTAT 1868
|||||
Db 1684 CAAAGAGAAACCAGATATGTTCAGACAAGAGAAATGTCAATCCTGTTTTCTGTATTGTAT 1743
QY 1869 GAGAACCAGAGCTGTGACTTCACAGAGAAATATACAACGCTTCTCCCAATCCAGCTGGA 1928
|||||
Db 1744 GAGAACCAGAGCTGTGACTTCACAGAGAAATATACAACGCTTCTCCCAATCCAGCTGGA 1803
QY 1929 GTGAGCTTGAGGATCCAGAGTTCCAAAGCCTCCAAACATCATGCAAGCATCAATGGCTAT 1988
|||||
Db 1804 GTGAGCTTGAGGATCCAGAGTTCCAAAGCCTCCAAACATCATGCAAGCATCAATGGCTAT 1863
QY 1989 GTTTTGTAGTTTGTGAGTTGTCAAGTTGTTGTCATGAGGTGGCATACTGTTACATTTCTA 2048
|||||
Db 1864 GTTTTGTAGTTTGTGAGTTGTCAAGTTGTTGTCATGAGGTGGCATACTGTTACATTTCTA 1923
QY 2049 AGCATTTGAGCAGACAGACTGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2108
|||||
Db 1924 AGCATTTGAGCAGACAGACTGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1983
QY 2109 AAAATGCTATGAAGACACACTACCCCTATTTCCATTTCTCAGAGAAACTGTCTTCAATG 2168
|||||
Db 1984 AAAATGCTATGAAGACACACTACCCCTATTTCCATTTCTCAGAGAAACTGTCTTCAATG 2043
QY 2169 TCGATGAAAAACCCAGGCTATGATTTCTGGGTGCGCACTCAGACTTTGGAACAGA 2228
|||||
Db 2044 TCGATGAAAAACCCAGGCTATGATTTCTGGGTGCGCACTCAGACTTTGGAACAGA 2103
QY 2229 GGCATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAAGAACACTGTTGATTATTACGAG 2288
|||||
Db 2104 GGCATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAAGAACACTGTTGATTATTACGAG 2163
QY 2289 GACAGTTATGAAGATTTTCAGCATACTGCTGAGTAAAAACAATGCCATTGAACCAAGA 2348
|||||
Db 2164 GACAGTTATGAAGATTTTCAGCATACTGCTGAGTAAAAACAATGCCATTGAACCAAGA 2223
QY 2349 AGCTTCTCCAGAACCCACAGTCTTGAACGCCCATCAACGGGAATACTGTTACTACT 2408
|||||
Db 2224 -----GAAATAACTGTTACTACT 2241
QY 2409 CTTCAGTCAGATCAAGAGAAATGCACTATGATGATACCATATCAGTTGAAATGAAGAAG 2468
|||||
Db 2242 CTTCAGTCAGATCAAGAGAAATGCACTATGATGATACCATATCAGTTGAAATGAAGAAG 2301
QY 2469 GAAGATTTTGACATTTATGATGAGAGTAAAAATCAGAGCCCCCGCAGCTTTCAAAAAGAAA 2528
|||||
Db 2302 GAAGATTTTGACATTTATGATGAGAGTAAAAATCAGAGCCCCCGCAGCTTTCAAAAAGAAA 2361

QY 2529 ACACGACACTATTTTATTTGCTGACAGTGGAGAGAGGCTCTGGGATTATGGGATGAGTACCTCC 2588
|||||
Db 2362 ACACGACACTATTTTATTTGCTGACAGTGGAGAGAGGCTCTGGGATTATGGGATGAGTACCTCC 2421
QY 2589 CCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTT 2648
|||||
Db 2422 CCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTT 2481
QY 2649 TTCCAGGAATTTACTGATGGCTCCTTTACTCAGCCCTTATACCCTGGAGAACTAAATGAA 2708
|||||
Db 2482 TTCCAGGAATTTACTGATGGCTCCTTTACTCAGCCCTTATACCCTGGAGAACTAAATGAA 2541
QY 2709 CATTTGGGACTCCTGGGGCCATATATAGAGCAGAAAGTTGAAGATATATCATGTGTAAT 2768
|||||
Db 2542 CATTTGGGACTCCTGGGGCCATATATAGAGCAGAAAGTTGAAGATATATCATGTGTAAT 2601
QY 2769 TTCAAGAAATCAGGCTCTCGTCCCTATTCCTTATTTCTAGCCCTTATTTCTTATGAGGAA 2828
|||||
Db 2602 TTCAAGAAATCAGGCTCTCGTCCCTATTCCTTATTTCTAGCCCTTATTTCTTATGAGGAA 2661
QY 2829 GATCAGAGGCAAGGAGCAGAACTTAGAAAAAATTTGTCAAGCCTAATGAAACCAAAACT 2888
|||||
Db 2662 GATCAGAGGCAAGGAGCAGAACTTAGAAAAAATTTGTCAAGCCTAATGAAACCAAAACT 2721
QY 2889 TACTTTTGGAAAGTGCACATCATATGCGACCCACTAAAGATGAGTTTGACTGCCAAAGCC 2948
|||||
Db 2722 TACTTTTGGAAAGTGCACATCATATGCGACCCACTAAAGATGAGTTTGACTGCCAAAGCC 2781
QY 2949 TGGGCTTATTTCTCTGATGTTGACCTGGAAAAAGATGTGCACCTCAGGCTGATTTGGACCC 3008
|||||
Db 2782 TGGGCTTATTTCTCTGATGTTGACCTGGAAAAAGATGTGCACCTCAGGCTGATTTGGACCC 2841
QY 3009 CTTCGTGCTGCCACACTAACACACTGAACCCCTGCTCATGGAGACAAGTGAAGTACAG 3068
|||||
Db 2842 CTTCGTGCTGCCACACTAACACACTGAACCCCTGCTCATGGAGACAAGTGAAGTACAG 2901
QY 3069 GAAATTTGCTCTGTTTTTACCACTTTTGATGAGAACCAAAAGCTGTACTCTGATAAAT 3128
|||||
Db 2902 GAAATTTGCTCTGTTTTTACCACTTTTGATGAGAACCAAAAGCTGTACTCTGATAAAT 2961
QY 3129 ATGGAAGAAGAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAG 3188
|||||
Db 2962 ATGGAAGAAGAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAG 3021
QY 3189 AATTATGCTTCCATGCAATCAATGGCTACATATGATATGATACACTACCTGGCTTAGTAATG 3248
|||||
Db 3022 AATTATGCTTCCATGCAATCAATGGCTACATATGATATGATATGATACACTACCTGGCTTAGTAATG 3081
QY 3249 GCTCAGGATCAAAAGGATTCGATGCTATCTGCTCAGCATGGGCGACCAATGAAAACATCCAT 3308
|||||
Db 3082 GCTCAGGATCAAAAGGATTCGATGCTATCTGCTCAGCATGGGCGACCAATGAAAACATCCAT 3141
QY 3309 TCTATTCAATTCAGTGGACATGTGTTCACTGTGACGAAAAAAGAGAGATATAAATGGCA 3368
|||||
Db 3142 TCTATTCAATTCAGTGGACATGTGTTCACTGTGACGAAAAAAGAGAGATATAAATGGCA 3201
QY 3369 CTGTACAATCTCTATCCAGGTGTTTTTGAGACAGTGAATGTTTACCATCCAAAGCTGA 3428
|||||
Db 3202 CTGTACAATCTCTATCCAGGTGTTTTTGAGACAGTGAATGTTTACCATCCAAAGCTGA 3261
QY 3429 ATTTGGCGGGTGAATGCTTATTTGGCGGACATCTACATGCTGGGATGAGCACACTTTT 3488
|||||
Db 3262 ATTTGGCGGGTGAATGCTTATTTGGCGGACATCTACATGCTGGGATGAGCACACTTTT 3321
QY 3489 CTGTTGACAGCAATTAAGTGTACAGCTCCCTGGGAATGGCTTCTGGACACATTAGAGAT 3548
|||||
Db 3322 CTGTTGACAGCAATTAAGTGTACAGCTCCCTGGGAATGGCTTCTGGACACATTAGAGAT 3381
QY 3549 TTTTCAGATTACAGCTTCAGAGCAATATGACAGTGGGCCCCAAAGCTGGCCAGACTTCAT 3608
|||||
Db 3382 TTTTCAGATTACAGCTTCAGAGCAATATGACAGTGGGCCCCAAAGCTGGCCAGACTTCAT 3441
QY 3609 TATTCCGATCAATCAATGCTTGAGACCAAGAGAGCCCTTTTCTTGATCAAGGTGAT 3668

```
|||||
Db 3442 TATTCGGATCAATCAATGCGCTGAGAGCAAGAGCCCTTTCTTGATCAAGTGGAAT 3501
QY 3669 CTGTTGGCACCACATGATTTATTCACGGCATCAAGACCGAGGTGCCCGTCAGAGTCTCC 3728
Db 3502 CTGTTGGCACCACATGATTTATTCACGGCATCAAGACCGAGGTGCCCGTCAGAGTCTCC 3561
QY 3729 AGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGAGAAGTGGCAGACT 3788
Db 3562 AGCCTCTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGAGAAGTGGCAGACT 3621
QY 3789 TATCGAGGAATTCACATGGAACCTTAATGCTCTCTTTGGCAATGTGATTCATCTGGG 3848
Db 3622 TATCGAGGAATTCACATGGAACCTTAATGCTCTCTTTGGCAATGTGATTCATCTGGG 3681
QY 3849 ATAAACACATATTTTAAACCTCCATATATGCTGATACATCCGTTTGCACCCACT 3908
Db 3682 ATAAACACATATTTTAAACCTCCATATATGCTGATACATCCGTTTGCACCCACT 3741
QY 3909 CATATAGCATTCGACGACTCTTCGATGGAGTGTGATGGGCTGTGATTTAATAGTGC 3968
3742 CATATAGCATTCGACGACTCTTCGATGGAGTGTGATGGGCTGTGATTTAATAGTGC 3801
QY 3969 AGCATGCCATTTGGAGTGGAGAGTAAGCAATATCAGATGCACAGATTACTGCTTATCC 4028
Db 3802 AGCATGCCATTTGGAGTGGAGAGTAAGCAATATCAGATGCACAGATTACTGCTTATCC 3861
QY 4029 TACTTTACCAATATGTTTGGCACCCTGGTCTCTTCAAAAAGCTGCACCTCCAAAGG 4088
Db 3862 TACTTTACCAATATGTTTGGCACCCTGGTCTCTTCAAAAAGCTGCACCTCCAAAGG 3921
QY 4089 AGGAGTAATGCGCTGGAGACCTCAGGTGATTAATCCAAAAGAGTGGCTGCAAGTGGACTTC 4148
Db 3922 AGGAGTAATGCGCTGGAGACCTCAGGTGATTAATCCAAAAGAGTGGCTGCAAGTGGACTTC 3981
QY 4149 CAGAAGCAATGAAAGTTCACAGGAGTAACTACTCAGGAGTAAATCTCTGTACCAAC 4208
Db 3982 CAGAAGCAATGAAAGTTCACAGGAGTAACTACTCAGGAGTAAATCTCTGTACCAAC 4041
QY 4209 ATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTCTT 4268
Db 4042 ATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTCTT 4101
QY 4269 TTTCAAGATGGCAAGTAAAGTTTTCAGGAAATCAAGACTCTTCAACACTGTGTG 4328
Db 4102 TTTCAAGATGGCAAGTAAAGTTTTCAGGAAATCAAGACTCTTCAACACTGTGTG 4161
QY 4329 AACTCTCTAGACCCACCGTACTGACTGCTACCTTGAATTCAACCCCAAGAGTGGGTG 4388
4162 AACTCTCTAGACCCACCGTACTGACTGCTACCTTGAATTCAACCCCAAGAGTGGGTG 4221
QY 4389 CACCAAGATGGCCCTGAGGATGGAGTTCGTGGGCTGCGAGGCACAGAGCTCTAC 4442
Db 4222 CACCAAGATGGCCCTGAGGATGGAGTTCGTGGGCTGCGAGGCACAGAGCTCTAC 4275
```

RESULT 9

```
N80447
N80447 standard; DNA; 4272 BP.
N80447:
10-OCT-1990 (first entry)
DE Modified factor VIII:C sequence with the R740-D1658 deletion.
KW Modified factor VIII:C; haemophilia; procoagulant;
OS Homo sapiens.
PN WO800831-A.
PU 11-FEB-1988.
PR 31-JUL-1987; U01814.
PA (BioJ) Biogen NV (PASE/).
PT PASEK MP.
NP WPI: 88-049866/07.
DR P-PSDB: P80268.
```

PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA
PT encoding maturation polypeptide, useful for high yield transformation.
PS Claim 3; Page 47-48-49-50; 97pp; English.
CC The RD deletion removes the DNA from Ser 741 to Ser 1657. A major part
CC of the sequence encoding the maturation polypeptide of
CC factor VIII:C is deleted, i.e. Gln 744 - Asp 1563.
CC The full length Factor VIII:C cDNA has two changes with respect to the
CC published sequence (EPO application 160457):
CC CTA to CTA at Leu 242 and TTC to CTC change at amino acid residue 1880
CC (Phe to Leu). The product is produced in approx. 20 times higher
CC yields than previous recombinant produced factor VIII:C and are more
CC easily purified. The peptide is used for treating haemophilia A, both
CC acute and prolonged bleeding.
CC See also N80444 and N80446.
SQ Sequence 4272 BP; 1243 A; 941 C; 944 G; 1144 T;

Query Match 87.2%; Score 4213.6; DB 1; Length 4272;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 4268; Conservative 0; Mismatches 4; Indels 42; Gaps 1;

```
QY 129 GCCACCAAGATACTACTCTGGGTGCACTGGAACCTGTCATGGACTATATGCAGAGTAT 188
Db 1 GCCACCAAGATACTACTCTGGGTGCACTGGAACCTGTCATGGACTATATGCAGAGTAT 60
QY 189 CTCGGTGAAGCTGCGCTGGAGCGCAAGATTCTCTCCATAGAGTGCCAAAATCTTTCCATTC 248
Db 61 CTCGGTGAAGCTGCGCTGGAGCGCAAGATTCTCTCCATAGAGTGCCAAAATCTTTCCATTC 120
QY 249 AACACCTCAGTCTGTATCAAAAAGACTCTGTTGTAGAAATTCACGGATCACCTTTTCAAC 308
Db 121 AACACCTCAGTCTGTATCAAAAAGACTCTGTTGTAGAAATTCACGGATCACCTTTTCAAC 180
QY 309 ATGCGTAAGCCCAAGGCCACCCTGGATGGGTCTGTAGTCTTACCATCCAGGCTGAGGTT 368
Db 181 ATGCGTAAGCCCAAGGCCACCCTGGATGGGTCTGTAGTCTTACCATCCAGGCTGAGGTT 240
QY 369 TATGATACAGTGTGTCATTAACATTAAGACATGGCTCCATCCCTGTCAGTCTTCAATGCT 428
Db 241 TATGATACAGTGTGTCATTAACATTAAGACATGGCTCCATCCCTGTCAGTCTTCAATGCT 300
QY 429 GTTGGTATCTCTACTGAAAGCTTCTGAGGAGCTGAATATGATGATCAGACAGTCAA 488
Db 301 GTTGGTATCTCTACTGAAAGCTTCTGAGGAGCTGAATATGATGATCAGACAGTCAA 360
QY 489 AGGAGAAAGAGATGATTAAGTCTTCCCTGGTGAAGCCATACATATGCTGAGAGTC 548
Db 361 AGGAGAAAGAGATGATTAAGTCTTCCCTGGTGAAGCCATACATATGCTGAGAGTC 420
QY 549 CTGAAGAGAAATGATGATGAGGCTGACCCACAGTGCCTTACCTACTCATATCTTCT 608
Db 421 CTGAAGAGAAATGATGATGAGGCTGACCCACAGTGCCTTACCTACTCATATCTTCT 480
QY 609 CATGTGACCTGTAAAGACTTGAATTCAAGGCTCATTTGAGCCCTACTAGTATGAGA 668
Db 481 CATGTGACCTGTAAAGACTTGAATTCAAGGCTCATTTGAGCCCTACTAGTATGAGA 540
QY 669 GAAGGAGTCTGGCAAGGAAAGACACAGACCTTGCAAAATTTACTACTTTTCT 728
Db 541 GAAGGAGTCTGGCAAGGAAAGACACACACCTTGCAAAATTTACTACTTTTCT 600
QY 729 GTATTTGATGAAGGAAAGTTGCACTCAGAAACAAGAACTCTTGTATGACAGATAGG 788
Db 601 GTATTTGATGAAGGAAAGTTGCACTCAGAAACAAGAACTCTTGTATGACAGATAGG 660
QY 789 GATGCTCATCTGCTCGGCGCTGGCTTAAATGCACACAGTCAATGTTATGTAACAGG 848
Db 661 GATGCTCATCTGCTCGGCGCTGGCTTAAATGCACACAGTCAATGTTATGTAACAGG 720
QY 849 TCTCTGCCAGTCTGATTTGGATGCCACAGGAATCAAGTCTATTTGGCATGTGATGGAATG 908
Db 721 TCTCTGCCAGTCTGATTTGGATGCCACAGGAATCAAGTCTATTTGGCATGTGATGGAATG 780
```

QY	909	GGCACCACCTCCTGAAGTGCACCTCAATATCTCTCGAAGGTCCACACATTTCTGTGAGAAC	968
Db	781	GGCACCACCTCCTGAAGTGCACCTCAATATCTCTCGAAGGTCCACACATTTCTGTGAGAAC	840
QY	969	CATGCCACGCGCTCTTGGAAATCTCGCCAAATAACTTTCTTACTGCTCAAAACACTCTTG	1028
Db	841	CATGCCACGCGCTCTTGGAAATCTCGCCAAATAACTTTCTTACTGCTCAAAACACTCTTG	900
QY	1029	ATGACCTTGGACAGTTTCTACTGTGTTTGTATATCTTCTCCACCAACATGATGGCATG	1088
Db	901	ATGACCTTGGACAGTTTCTACTGTGTTTGTATATCTTCTCCACCAACATGATGGCATG	960
QY	1089	GAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGAAACCCCACTACGAATGAAAAATAAT	1148
Db	961	GAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGAAACCCCACTACGAATGAAAAATAAT	1020
QY	1149	GAAGAAGCGGGAAGACTATGATGATGATCTTACTGATCTTGAATGGAATGTGTACGCTTT	1209
Db	1021	GAAGAAGCGGGAAGACTATGATGATGATCTTACTGATCTTGAATGGAATGTGTACGCTTT	1080
QY	1209	GATGATGACAACTCTCCTCTCTTATCCAAATTCGTCAGTTGCCAAGAAGCATCCTAAA	1268
Db	1081	GATGATGACAACTCTCCTCTCTTATCCAAATTCGTCAGTTGCCAAGAAGCATCCTAAA	1140
QY	1269	ACTTGGGTACATTACATTTGCTGCTGAGAGAGAGACTGGACTATGCTCCCTTAGCTCTC	1328
Db	1141	ACTTGGGTACATTACATTTGCTGCTGAGAGAGAGACTGGACTATGCTCCCTTAGCTCTC	1200
QY	1329	GCCCCCGATGACAGAACTTATAAAAGTCATATTTGAAACAATGGCCCTCAGCGGATTGGT	1388
Db	1201	GCCCCCGATGACAGAACTTATAAAAGTCATATTTGAAACAATGGCCCTCAGCGGATTGGT	1260
QY	1389	AGGAAGTACAAAAAAGTCCGATTATATGECATACACAGATGAACCTTTAAGACTCGTGA	1448
Db	1261	AGGAAGTACAAAAAAGTCCGATTATATGECATACACAGATGAACCTTTAAGACTCGTGA	1320
QY	1449	GCTATTCAGCATGAATTCAGGAATCTTGGGACCTTTACTTTATGGGGAAGTTGAGACACA	1508
Db	1321	GCTATTCAGCATGAATTCAGGAATCTTGGGACCTTTACTTTATGGGGAAGTTGAGACACA	1380
QY	1509	CTGTGATTTATATTTAAGAAATCAAGCAAGCAGACCATATTAACATCTACCCCTCAGGAATC	1568
Db	1381	CTGTGATTTATATTTAAGAAATCAAGCAAGCAGACCATATTAACATCTACCCCTCAGGAATC	1440
QY	1569	ACTGATGTCCGTCCTTTGTATTCAAGGAGATTACCAAAAGGTGTAAAACATTTGAAAGAT	1628
Db	1441	ACTGATGTCCGTCCTTTGTATTCAAGGAGATTACCAAAAGGTGTAAAACATTTGAAAGAT	1500
QY	1629	TTTCCAAATCTGCCCAGAGAAATATTTCAAAATATTAATGAGACAGTACTGTAGAAAGTGG	1688
Db	1501	TTTCCAAATCTGCCCAGAGAAATATTTCAAAATATTAATGAGACAGTACTGTAGAAAGTGG	1560
QY	1689	CCAACCTAAATCAGATCCTCGGTGCCGTGACCCGCTATTACTCTAGTTTCGTTAATATGAG	1748
Db	1561	CCAACCTAAATCAGATCCTCGGTGCCGTGACCCGCTATTACTCTAGTTTCGTTAATATGAG	1620
QY	1749	AGAGATCTAGCTTACAGGACTCATATGGCCCCCTCTCTCATCTGCTACAAAGAAATCTGTAGAT	1808
Db	1621	AGAGATCTAGCTTACAGGACTCATATGGCCCCCTCTCTCATCTGCTACAAAGAAATCTGTAGAT	1680
QY	1809	CAAGAGGAAACCCAGATATATGTGACACAAAGAGGAATGTCAATCTGTTTCTGTATTGGAT	1868
Db	1681	CAAGAGGAAACCCAGATATATGTGACACAAAGAGGAATGTCAATCTGTTTCTGTATTGGAT	1740
QY	1869	GAGAACCAGAGCTGTACCTCAGAGAGAAATATACAAAGCTTTCTCCCCAATCCAGCTGGA	1928
Db	1741	GAGAACCAGAGCTGTACCTCAGAGAGAAATATACAAAGCTTTCTCCCCAATCCAGCTGGA	1800
QY	1929	GTCGACCTTGAGGATCCAGAGTTCCAAAGCCCTCCAAACATCATGCAAGCATCAATGGCTAT	1988
Db	1801	GTCGACCTTGAGGATCCAGAGTTCCAAAGCCCTCCAAACATCATGCAAGCATCAATGGCTAT	1860
QY	1989	GTTTTTGATAGTTTGCAGTTGTCTAGTTTGTTCATGAGGTTGGCATACTGTGATCTTCTA	2048

Db	1861	GTTTTGAATAGTTTGACAGTTGTACAGTTTGTGTCATGAGAGGTGGCACTACTGGTACTGTA	1920
QY	2049	AGCATTGGAGCACAGACTGACCTTCTTCTGTCTTCTCTCTGTGATATACCTTCAAAAC	2108
Db	1921	AGCATTGGAGCACAGACTGACCTTCTTCTGTCTTCTCTGTGATATACCTTCAAAAC	1980
QY	2109	AAATGCTATGAAGACACACTCACCTATTTCCATTTCTCAGAGAAACTGTCTCATG	2166
Db	1981	AAATGCTATGAAGACACACTCACCTATTTCCATTTCTCAGAGAAACTGTCTCATG	2040
QY	2169	TCGATGAAAAACCAGGCTCTATGATTTCTGGGGTGGCCAACTCAGACTTTCGAAACGA	2228
Db	2041	TCGATGAAAAACCAGGCTCTATGATTTCTGGGGTGGCCAACTCAGACTTTCGAAACGA	2100
QY	2229	GGCATGACCGCTTACTGAAAGTTTCTAGTTGTGACAAAGAACACTGGTGATTTATACGAG	2288
Db	2101	GGCATGACCGCTTACTGAAAGTTTCTAGTTGTGACAAAGAACACTGGTGATTTATACGAG	2160
QY	2289	GACAGTTATGAAGATATTTCCAGCATCTTGTGAGTAAAAACAATGCCATTGAACCAAGA	2348
Db	2161	GACAGTTATGAAGATATTTCCAGCATCTTGTGAGTAAAAACAATGCCATTGAACCAAGA	2220
QY	2349	AGCTTCTCCAGAACCCACAGTCTTGAAACGCCATCAACGGGAATACTCGTACTACT	2408
Db	2221	-----GAAATACTCGTACTACT	2238
QY	2409	CTTCAGTCAGATCAAGAGAAATTGACATATGATGATACCATATCAGTTGAATGAAGAG	2468
Db	2239	CTTCAGTCAGATCAAGAGAAATTGACATATGATGATACCATATCAGTTGAATGAAGAG	2298
QY	2469	GAAGATTTTGACATTTATGATGAGGATGAAAAATCAGAGCCCCCGCAGCTTTCAAAAGAA	2528
Db	2299	GAAGATTTTGACATTTATGATGAGGATGAAAAATCAGAGCCCCCGCAGCTTTCAAAAGAA	2358
QY	2529	ACACGACACTATTTTATGCTGACGTGAGAGGCTCTGGATATGGGATGAGTAGCTCC	2588
Db	2359	ACACGACACTATTTTATGCTGACGTGAGAGGCTCTGGATATGGGATGAGTAGCTCC	2418
QY	2589	CCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCACAGAAAGTGT	2648
Db	2419	CCACATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCACAGAAAGTGT	2478
QY	2649	TTCCAGGAATTTACTGATGGCTCCCTTACTCAGCCCTTATACCCTGAGAACTAAATGAA	2708
Db	2479	TTCCAGGAATTTACTGATGGCTCCCTTACTCAGCCCTTATACCCTGAGAACTAAATGAA	2538
QY	2709	CATTGGGACTCCTGGGGCCATATATAAGAGCAGAAAGTTGAAGATAATATCATGTAACT	2768
Db	2539	CATTGGGACTCCTGGGGCCATATATAAGAGCAGAAAGTTGAAGATAATATCATGTAACT	2598
QY	2769	TTCAGAAATCAGGCTCTCGTCCCTATTTCTCTATTTCTAGCCTTATTTCTTATGAGAA	2828
Db	2599	TTCAGAAATCAGGCTCTCGTCCCTATTTCTCTATTTCTAGCCTTATTTCTTATGAGAA	2658
QY	2829	GATCAGAGCAAGGACAGAACTAGAAAAAACTTTGTCAAGCCTAATGAACCAAAACT	2888
Db	2659	GATCAGAGCAAGGACAGAACTAGAAAAAACTTTGTCAAGCCTAATGAACCAAAACT	2718
QY	2889	TACTTTGGAAAGTGCAACATCATATGCGACCCACTAAAGATGAGTTTGACTGCAAGCC	2948
Db	2719	TACTTTGGAAAGTGCAACATCATATGCGACCCACTAAAGATGAGTTTGACTGCAAGCC	2778
QY	2949	TGGGCTATTTCTGATGTTGACCTGAAAAAAGATGTGCACTCAGGCTGATTTGGACCC	3008
Db	2779	TGGGCTATTTCTGATGTTGACCTGAAAAAAGATGTGCACTCAGGCTGATTTGGACCC	2838
QY	3009	CTTCGCTGTCACACTAACAACACTGAACCTGCTCATGGGAGACAAGTGCAGTACAG	3068
Db	2839	CTTCGCTGTCACACTAACAACACTGAACCTGCTCATGGGAGACAAGTGCAGTACAG	2898
QY	3069	GAATTTGCTCTGTTTTTACCACTCTTGATGAGACCAAAAGCTGTACTTCACTGAAT	3128

Db 2899 GAATTTGCTCTGTTTTTTCACCAATCTTGATGAGACCAAAAGCTGTACTTCACTGAAAT 2958
QY 3129 ATGGAAGAAACATGCGAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAG 3188
Db 2959 ATGGAAGAAACATGCGAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAG 3018
QY 3189 AATTATCGCTTCATGCAATCAATGCTACATATGATACACTACTGCTGCTTAATG 3248
Db 3019 AATTATCGCTTCATGCAATCAATGCTACATATGATACACTACTGCTGCTTAATG 3078
QY 3249 GCTCAGGATCAAGGATTCGATGATATCTGCTCAGCATGGGACCAATGAAAAATCCAT 3308
Db 3079 GCTCAGGATCAAGGATTCGATGATATCTGCTCAGCATGGGACCAATGAAAAATCCAT 3138
QY 3309 TCTATTCAATTCAGTGACATGTTCTCACTGTACGAAAAAAGAGAGTATATAATGCA 3368
Db 3139 TCTATTCAATTCAGTGACATGTTCTCACTGTACGAAAAAAGAGAGTATATAATGCA 3198
QY 3369 CTGTACAATCTCTATCCAGGTGTTTTTGGAGACAGTGAATGTACCAATCCAAAGCTGGA 3428
Db 3199 CTGTACAATCTCTATCCAGGTGTTTTTGGAGACAGTGAATGTACCAATCCAAAGCTGGA 3258
QY 3429 AATTGGCGGGTGAATGCCCTTATTTGGCGAGCATCTACATGCTGGGATGAGCACACTTTT 3488
Db 3259 AATTGGCGGGTGAATGCCCTTATTTGGCGAGCATCTACATGCTGGGATGAGCACACTTTT 3318
QY 3489 CTGGTGTACAGCAATAAGTGTACAGACTCCCTGGGAATGGCTTGTGACACATAGAGAT 3548
Db 3319 CTGGTGTACAGCAATAAGTGTACAGACTCCCTGGGAATGGCTTGTGACACATAGAGAT 3378
QY 3549 TTTCAGATTACAGCTTCAGAGCAATATGAGACAGTGGGCCCAAGCTGGCCAGACTTCAT 3608
Db 3379 TTTCAGATTACAGCTTCAGAGCAATATGAGACAGTGGGCCCAAGCTGGCCAGACTTCAT 3438
QY 3609 TATTCGCGATCAATCAATGCCCTGAGCACCAAGAGCCCTTTCTTGATCAAGTGGAT 3668
Db 3439 TATTCGCGATCAATCAATGCCCTGAGCACCAAGAGCCCTTTCTTGATCAAGTGGAT 3498
QY 3669 CTGTGGCACCAATGATTATTCAGCGGCATCAAGACCCAGGGTGGCCCTCAGAACTTCTCC 3728
Db 3499 CTGTGGCACCAATGATTATTCAGCGGCATCAAGACCCAGGGTGGCCCTCAGAACTTCTCC 3558
QY 3729 AGCCTTACATCTCTCAGTTTATCATCATATAGTCTTGATGGGAAGAAGTGGAGACT 3788
Db 3559 AGCCTTACATCTCTCAGTTTATCATCATATAGTCTTGATGGGAAGAAGTGGAGACT 3618
QY 3789 TATCGAGGAATTCACATGGAACCTTAATGCTCTTTGGCAATGTGGATTCATCTGGG 3848
Db 3619 TATCGAGGAATTCACATGGAACCTTAATGCTCTTTGGCAATGTGGATTCATCTGGG 3678
QY 3849 ATAAACACAAATATTTTAAACCCCAATATATGCTCGATACATCCGTTTGCACCCCACT 3908
Db 3679 ATAAACACAAATATTTTAAACCCCAATATATGCTCGATACATCCGTTTGCACCCCACT 3738
QY 3909 CATATAGCATTCGACGACACTCTTCGCATGAGTGTGATGGGCTGTGATTTAAATAGTGC 3968
Db 3739 CATATAGCATTCGACGACACTCTTCGCATGAGTGTGATGGGCTGTGATTTAAATAGTGC 3798
QY 3969 AGCATGCCATTTGGGAATGAGAGTAAGCAATATCAGATGCACAGATTACTGCTTCATCC 4028
Db 3799 AGCATGCCATTTGGGAATGAGAGTAAGCAATATCAGATGCACAGATTACTGCTTCATCC 3858
QY 4029 TACTTTACCAATATGTTTGGCACCCTGGTCTCTTCAAAAAGCTCGACTTCACCTCCAAGG 4088
Db 3859 TACTTTACCAATATGTTTGGCACCCTGGTCTCTTCAAAAAGCTCGACTTCACCTCCAAGG 3918
QY 4089 AGGAGTAATGCCGTGAGAGACCTCAGGTGAATATCCAAAAAGAGTGGCTGCAAGTGAAGTTC 4148
Db 3919 AGGAGTAATGCCGTGAGAGACCTCAGGTGAATATCCAAAAAGAGTGGCTGCAAGTGAAGTTC 3978
QY 4149 CAGAGACAAATGAAGTCAAGAGTAATCTACTCAGGAGTAATAATCTCTGCTTACCAGC 4208
Db 3979 CAGAGACAAATGAAGTCAAGAGTAATCTACTCAGGAGTAATAATCTCTGCTTACCAGC 4038

QY 4209 ATGTATGTAAGAGTCTCCTCATCTCCAGACATCAAGATGGCCATCAGTGCATCTCTT 4268
Db 4039 ATGTATGTAAGAGTCTCCTCATCTCCAGACATCAAGATGGCCATCAGTGCATCTCTT 4098
QY 4269 TTTCAGATGGCAAGTAAGGTTTTTCAGGGAATCAAGACTCCTTACACACTGTGTG 4328
Db 4099 TTTCAGATGGCAAGTAAGGTTTTTCAGGGAATCAAGACTCCTTACACACTGTGTG 4158
QY 4329 AACTCTAGACCCACCGTTACTGACTCGCTACCTTGAATTCACCCCAAGAGTGGTG 4388
Db 4159 AACTCTAGACCCACCGTTACTGACTCGCTACCTTGAATTCACCCCAAGAGTGGTG 4218
QY 4389 CACCAGATTGCCCTGAGGATGAGGTTCTGGGCTGGAGGCAAGGACCTCTAC 4442
Db 4219 CACCAGATTGCCCTGAGGATGAGGTTCTGGGCTGGAGGCAAGGACCTCTAC 4272

RESULT 10
N80444
ID N80444 standard; DNA; 4545 BP.
AC N80444;
DE 10-OCT-1990 (first entry)
KW Modified factor VIII:C sequence with the Q744-D1563 deletion.
KW blood coagulation; QD deletion; ss.
OS Homo sapiens.
PN W0800831-A.
PD 11-FEB-1988.
PF 31-JUL-1987; U01814.
PR 01-AUG-1986; US-893375.
PA (BIOU) Biogen NV (PASE/).
PI PASEK MP.
DR WPI; 88-049866/07.
DR P-PSDB; P80265.
PT New DNA sequences encoding modified factor VIII:C - with deletion of DNA
PT encoding maturation polypeptide, useful for high yield transformation.
PS Claim 3; Page 38-39-40-41; 97pp; English.
CC A major part of the sequence encoding the maturation polypeptide of
CC factor VIII:C is deleted, i.e. Gln 744 - Asp 1563. The QD deletion
CC retains approximately 90 amino acids of the maturation polypeptide
CC (four amino acids at the N-terminal end and 86 amino acids at
CC the C-terminal end). The full length Factor VIII:C cDNA has two
CC changes with respect to the published sequence (EPO application 160457):
CC CTG to CTA at Leu 242 and TTC to CTC change at amino acid residue 1880
CC (Phe to Leu). The product is produced in approx. 20 times higher
CC yields than previous recombinant produced factor VIII:C and are more
CC easily purified. The peptide is used for treating haemophilia A, both
CC acute and prolonged bleeding.
CC See also N80446 and N80447.
SQ Sequence 4545 BP; 1345 A; 1002 C; 1003 G; 1195 T;

Query Match 84.2%; Score 4069.6; DB 1; Length 4545;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 4310; Conservative 0; Mismatches 4; Indels 228; Gaps 1;

QY 129 GCCACCAGAACTACTACCTGGTGCAGTGAAGTGTATGGAAGTATATGCAAGTGTAT 188
Db 4 GCCACCAGAACTACTACCTGGTGCAGTGAAGTGTATGGAAGTATATGCAAGTGTAT 63
QY 189 CTCGGTGAGCTGCTGTGAGCGCAAGATTCTCTCTAGAGTGCACAAATCTTTCCATTC 248
Db 64 CTCGGTGAGCTGCTGTGAGCGCAAGATTCTCTCTAGAGTGCACAAATCTTTCCATTC 123
QY 249 AACACCTCAGTCTGTACAAAAGACTCTGTTGTAGATTCACGGATCACCCTTTCAAC 308
Db 124 AACACCTCAGTCTGTACAAAAGACTCTGTTGTAGATTCACGGATCACCCTTTCAAC 183
QY 309 ATCGCTAAGCCAGGCCACCTGGATGGTCTGCTAGTGTCTTACCATCCAGGCTGAGGT 368
Db 184 ATCGCTAAGCCAGGCCACCTGGATGGTCTGCTAGTGTCTTACCATCCAGGCTGAGGT 243

QY 369 TATGATACAGTGGTCATTTACACTTAAGACATGGCTTCCCATCTGTCAGTCTTTCATGCT 428
|||||
Db 244 TATGATACAGTGGTCATTTACACTTAAGACATGGCTTCCCATCTGTCAGTCTTTCATGCT 303
QY 429 GTTGGTGTATCTTACTGGAAGCTTCTGAGGGAGCTGAATATGATGATCAGACAGCTCAA 488
|||||
Db 304 GTTGGTGTATCTTACTGGAAGCTTCTGAGGGAGCTGAATATGATGATCAGACAGCTCAA 363
QY 489 AGGAGAGAGAGATGATTAAGTCTTCCCTGGTGGAGGCCATATATGCTGGCAGCTC 548
|||||
Db 364 AGGAGAGAGAGATGATTAAGTCTTCCCTGGTGGAGGCCATATATGCTGGCAGCTC 423
QY 549 CTGAAGAGAGATGTCATAGGCTCTGACCCACTGTGCTTACTACTATCTTCT 608
|||||
Db 424 CTGAAGAGAGATGTCATAGGCTCTGACCCACTGTGCTTACTACTATCTTCT 483
QY 609 CATGTGACCTGTGTAAGACTTGAATTCAGGCCCTCATTTGAGGCCCTACTAGTATGTA 668
|||||
Db 484 CATGTGACCTGTGTAAGACTTGAATTCAGGCCCTCATTTGAGGCCCTACTAGTATGTA 543
QY 669 GAAGGAGCTCTGGCCAAAGAAAGACACAGACCTTGCACAAATTTATCTACTTTTCT 728
|||||
Db 544 GAAGGAGCTCTGGCCAAAGAAAGACACACACCTTGCACAAATTTATCTACTTTTCT 603
QY 729 GTATTGATGAAGGAAAGTTGGCAGCTCAGAAACAAGAACTCTTGATGACAGATAGG 788
|||||
Db 604 GTATTGATGAAGGAAAGTTGGCAGCTCAGAAACAAGAACTCTTGATGACAGATAGG 663
QY 789 GATGCTGCATCTGCTGGGGCTGGCCCTAAATGCACACAGTCAATGGTTATGTAACAGG 848
|||||
Db 664 GATGCTGCATCTGCTGGGGCTGGCCCTAAATGCACACAGTCAATGGTTATGTAACAGG 723
QY 849 TCTTCGCCAGGCTGATTTGGATGGCCACAGGAAATCAGTCTATTTGGATGATGGAATG 908
|||||
Db 724 TCTTCGCCAGGCTGATTTGGATGGCCACAGGAAATCAGTCTATTTGGATGATGGAATG 783
QY 909 GGCACCACTCTGAAAGTGCACCTCAATATCTCTGAAAGTGCACACATTTCTTGTGAGAAC 968
|||||
Db 784 GGCACCACTCTGAAAGTGCACCTCAATATCTCTGAAAGTGCACACATTTCTTGTGAGAAC 843
QY 969 CATGGCCAGGCGTCTTGGAAATCTGCCCAATTAACCTTCTTACTGCTCAAACACACTCTG 1028
|||||
Db 844 CATGGCCAGGCGTCTTGGAAATCTGCCCAATTAACCTTCTTACTGCTCAAACACACTCTG 903
QY 1029 ATGACCTTGGACAGTTTCTACTGTTTTGTCTATATCTTCCACCAACATGATGCGATG 1088
|||||
Db 904 ATGACCTTGGACAGTTTCTACTGTTTTGTCTATATCTTCCACCAACATGATGCGATG 963
QY 1089 GAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGGAAACCCCACTACGAATGAAAAATAT 1148
|||||
Db 964 GAAGCTTATGTCAAAAGTAGACAGCTGTCCAGAGGAAACCCCACTACGAATGAAAAATAT 1023
QY 1149 GAAGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAATGATGATGATGATG 1208
|||||
Db 1024 GAAGAAGCGGAAGACTATGATGATGATCTTACTGATTTCTGAATGATGATGATGATG 1083
QY 1209 GATGATGACAACTCTCTCTCTTATTCAAATTCGCTCAGTTGCCAAGAAGCATCTTAA 1268
|||||
Db 1084 GATGATGACAACTCTCTCTCTTATTCAAATTCGCTCAGTTGCCAAGAAGCATCTTAA 1143
QY 1269 ACTTGGGTACATTAATGCTGTGTAAGAGAGAGGAGCTGGACTATGCTCCCTTAGTCTC 1328
|||||
Db 1144 ACTTGGGTACATTAATGCTGTGTAAGAGAGAGGAGGAGCTGGACTATGCTCCCTTAGTCTC 1203
QY 1329 GCCCCCGATGACAGAGTTATTAAGTCAATATTTGAACAATGGCCCTCAGCGGATGGT 1388
|||||
Db 1204 GCCCCCGATGACAGAGTTATTAAGTCAATATTTGAACAATGGCCCTCAGCGGATGGT 1263
QY 1389 AGGAAGTACAAAAAGTCCGATTTATGCGATACAGATGAACCTTTAAGACTCGTGAA 1448
|||||
Db 1264 AGGAAGTACAAAAAGTCCGATTTATGCGATACAGATGAACCTTTAAGACTCGTGAA 1323
QY 1449 GCTATTTCAGCATGAATCAGGAATCTTGGACCTTACTTTATGGGGAAGTTGAGACACA 1508

Db 1324 GCTATTTCAGCATGAATCAGGAATCTTGGACCTTTACTTTATGCGGGAAGTTGGAGACACA 1383
QY 1509 CTGTTGATTATATTTAAGAAATCAAGCAAGCAGACATATTAACATCTACCCCTCAGGAATC 1568
|||||
Db 1384 CTGTTGATTATATTTAAGAAATCAAGCAAGCAGACATATTAACATCTACCCCTCAGGAATC 1443
QY 1569 ACTGATGTCCTCTCTTGTATTCAGGAGATTTACCAAAAGGTGTAAACATTTGAAGAT 1628
|||||
Db 1444 ACTGATGTCCTCTCTTGTATTCAGGAGATTTACCAAAAGGTGTAAACATTTGAAGAT 1503
QY 1629 TTTCCAATTTGCCAGGAGAGAAATATTCAAATATTAATGAGACAGTGAAGTGAAGATGG 1688
|||||
Db 1504 TTTCCAATTTGCCAGGAGAGAAATATTCAAATATTAATGAGACAGTGAAGATGG 1563
QY 1689 CCAACTAAATCAGATCTCGGTCCTGACCCGCTATTACTAGTTTCTGTTAATATGAG 1748
|||||
Db 1564 CCAACTAAATCAGATCTCGGTCCTGACCCGCTATTACTAGTTTCTGTTAATATGAG 1623
QY 1749 AGAGATCTAGCTTCAGGACTATTTGGCCCTCTCTCATCTGCTACAAAGAAATCTAGAT 1808
|||||
Db 1624 AGAGATCTAGCTTCAGGACTATTTGGCCCTCTCTCATCTGCTACAAAGAAATCTAGAT 1683
QY 1809 CAAAGAGAAACAGATTAATGTTCAGACAAGAGAAATGTATCTGTTTCTGTTATTTGAT 1868
|||||
Db 1684 CAAAGAGAAACAGATTAATGTTCAGACAAGAGAAATGTATCTGTTTCTGTTATTTGAT 1743
QY 1869 GAGAACCAGAGCTGTACTCTCAGAGAAATATCAACGCTTCTCCCAATCCAGCTGA 1928
|||||
Db 1744 GAGAACCAGAGCTGTACTCTCAGAGAAATATCAACGCTTCTCCCAATCCAGCTGA 1803
QY 1929 GTGCAGCTTGAGATCCAGAGTTCCAAAGCTCCAAACATATGCACAGCATCAATGCTAT 1988
|||||
Db 1804 GTGCAGCTTGAGATCCAGAGTTCCAAAGCTCCAAACATATGCACAGCATCAATGCTAT 1863
QY 1989 GTTTTGTATAGTTTGAGTTGTGATGTTGTTGTTGATGAGGTGGCATATGCTATCTA 2048
|||||
Db 1864 GTTTTGTATAGTTTGAGTTGTGATGTTGTTGTTGATGAGGTGGCATATGCTATCTA 1923
QY 2049 AGCATTTGAGACAGACTGACTTCTCTGCTCTTCTCTGATATATACCTTCAAAAC 2108
|||||
Db 1924 AGCATTTGAGACAGACTGACTTCTCTGCTCTTCTCTGATATATACCTTCAAAAC 1983
QY 2109 AAAATGCTATGAAGACACACTCACCCATTTCCCATTTCTCAGAGAAACTGCTTCATG 2168
|||||
Db 1984 AAAATGCTATGAAGACACACTCACCCATTTCCCATTTCTCAGAGAAACTGCTTCATG 2043
QY 2169 TCGATGAAAAACCCAGGTTATGATTTCTGGGTCGCACAACCTCAGACTTTCGGAACAGA 2228
|||||
Db 2044 TCGATGAAAAACCCAGGTTATGATTTCTGGGTCGCACAACCTCAGACTTTCGGAACAGA 2103
QY 2229 GGCATGACCGCTTACTGAAGTTTCTAGTTGTGACAAGAACTGGTGAATTATACGAG 2288
|||||
Db 2104 GGCATGACCGCTTACTGAAGTTTCTAGTTGTGACAAGAACTGGTGAATTATACGAG 2163
QY 2289 GACAGTTATGAAGATATTTCAAGCATACTTCTGAGTAAACAAATGCCATTTGAACCAAGA 2348
|||||
Db 2164 GACAGTTATGAAGATATTTCAAGCATACTTCTGAGTAAACAAATGCCATTTGAACCAAGA 2223
QY 2349 AGCTTCTCCAG----- 2361
|||||
Db 2224 AGCTTCTCCAGATCTCTTGTCTTGGGATTAACCACTATGCTACTCAGATACCAAAAGAA 2283
QY 2361 ----- 2361
Db 2284 GAGTGAATTCCTCAAGAGAACTCACCAAGAAAAACAGCTTTTAAGAAAAAGATACCAT 2343
QY 2361 ----- 2361
Db 2344 TTGTCCCTGAACGCTTGTGAAGCAATCATGCAATAGCAGCAATATAATGAGGACAAAAAT 2403
QY 2361 ----- 2361

Db 2404 AAGCCGAAATAGAGTCACCTGGGCAAGCAAGGTAGGACTGAAAGCCTGTGCTCTCA 2463
QY 2361 AACCCACAGTCTTGAAACGCCATCAACGGGAAATACTGCTACTCTTCAGTCAGAT 2420
Db 2464 AACCCACAGTCTTGAAACGCCATCAACGGGAAATACTGCTACTCTTCAGTCAGAT 2523
QY 2421 CAAGAGAAATTGACTATGATGATACCATATCAGTTGAAATGAAGAAGATTGAC 2480
Db 2524 CAAGAGAAATTGACTATGATGATACCATATCAGTTGAAATGAAGAAGATTGAC 2583
QY 2481 ATTTATGATGAGGATGAAATCAGAGCCCCCGCAGCTTCAAAAAGAAAACACGACACTAT 2540
Db 2584 ATTTATGATGAGGATGAAATCAGAGCCCCCGCAGCTTCAAAAAGAAAACACGACACTAT 2643
QY 2541 TTTATGCTCAGTGGAGAGGCTCTGGGATTATGGATGAGTAGCTCCCAATGTTCTA 2600
Db 2644 TTTATGCTCAGTGGAGAGGCTCTGGGATTATGGATGAGTAGCTCCCAATGTTCTA 2703
QY 2601 AGAAGAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTCCAGAAATT 2660
Db 2704 AGAAGAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTCCAGAAATT 2763
QY 2661 ACTGATGGCTCTTACTCAGCCCTTATACCCTGAGAACTAAATGAACATTTGGACTC 2720
Db 2764 ACTGATGGCTCTTACTCAGCCCTTATACCCTGAGAACTAAATGAACATTTGGACTC 2823
QY 2721 CTGGGGCATATATAGAGCAGAGTGAAGATATATCATGTGTAATTCAGAAATCAG 2780
Db 2824 CTGGGGCATATATAGAGCAGAGTGAAGATATATCATGTGTAATTCAGAAATCAG 2883
QY 2781 GCCCTGCTCCCTATTCCTTCTATTCAGCCTTATTCCTTATGAGAGAGTCAAGGCAA 2840
Db 2884 GCCCTGCTCCCTATTCCTTCTATTCAGCCTTATTCCTTATGAGAGAGTCAAGGCAA 2943
QY 2841 GGAGCAGAACTAGAAAAAACTTTGTCAAGCCTAATGAACCAGAACTTCTTTGAAA 2900
Db 2944 GGAGCAGAACTAGAAAAAACTTTGTCAAGCCTAATGAACCAGAACTTCTTTGAAA 3003
QY 2901 GTGCAACATCATATGCGCAACCCACTAAAGATGAGTTGACTGCAAGCCTGGCTTATTC 2960
Db 3004 GTGCAACATCATATGCGCAACCCACTAAAGATGAGTTGACTGCAAGCCTGGCTTATTC 3063
QY 2961 TCTGATGTTGACCTGAAAAAAGATGTGCACTCAGGCTGATTTGACCCCTCTGCTGC 3020
Db 3064 TCTGATGTTGACCTGAAAAAAGATGTGCACTCAGGCTGATTTGACCCCTCTGCTGC 3123
QY 3021 CACACTAACACACTGAACCCCTGCTCATGGGAGAGACAAGTACAGTACAGAAATTTGCTCG 3080
Db 3124 CACACTAACACACTGAACCCCTGCTCATGGGAGAGACAAGTACAGTACAGAAATTTGCTCG 3183
QY 3081 TTTTTCACCATCTTTGATGAGACCAAAAGCTGTACTTCACTGAAATATGAAAAAGAAC 3140
Db 3184 TTTTTCACCATCTTTGATGAGACCAAAAGCTGTACTTCACTGAAATATGAAAAAGAAC 3243
QY 3141 TGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAAATATCCGCTTC 3200
Db 3244 TGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAAATATCCGCTTC 3303
QY 3201 CATGCAATCAATGGCTACATAATGATACACTACCTGGCTTAGTAATGGCTCAGGATCAA 3260
Db 3304 CATGCAATCAATGGCTACATAATGATACACTACCTGGCTTAGTAATGGCTCAGGATCAA 3363
QY 3261 AGGATTCGATGTTATCTGCTCAGCATGGGCAAGCAATGAACATCCATCTATTCATTTTC 3320
Db 3364 AGGATTCGATGTTATCTGCTCAGCATGGGCAAGCAATGAACATCCATCTATTCATTTTC 3423
QY 3321 AGTGACATGTGTTCACTGTACGAAAAAAGAGAGATTAATAATGGCACTGTACAATCTC 3380
Db 3424 AGTGACATGTGTTCACTGTACGAAAAAAGAGAGATTAATAATGGCACTGTACAATCTC 3483
QY 3381 TATCCAGGTGTTTGTGACACAGTGAAGATGTTACCATCCAAAGCTGGAATTTGGCGGGTG 3440
Db 3484 TATCCAGGTGTTTGTGACACAGTGAAGATGTTACCATCCAAAGCTGGAATTTGGCGGGTG 3543

QY 3441 GAATGCCCTATTTGGCGAGCATCTACATGCTGGGATGAGCACACTTTTCTGCTGACAGC 3500
Db 3544 GAATGCCCTATTTGGCGAGCATCTACATGCTGGGATGAGCACACTTTTCTGCTGACAGC 3603
QY 3501 AATAAGTCTCAGACTCCCTTGGGAATGCTTCTTGACACATTTAGAGATTTTCAGATTACA 3560
Db 3604 AATAAGTCTCAGACTCCCTTGGGAATGCTTCTTGACACATTTAGAGATTTTCAGATTACA 3663
QY 3561 GCTTCAGACAATATGACAGAGTGGGCCCCAAAGCTGGCCAGACTTCATTATTCGGATCA 3620
Db 3664 GCTTCAGACAATATGACAGAGTGGGCCCCAAAGCTGGCCAGACTTCATTATTCGGATCA 3723
QY 3621 ATCAATGCCCTGAGCACCAAGAGGCCCTTTCTTGATCAAGGTGATCTGTGGCACCA 3680
Db 3724 ATCAATGCCCTGAGCACCAAGAGGCCCTTTCTTGATCAAGGTGATCTGTGGCACCA 3783
QY 3681 ATGATTAATTCAGGCAATCAAGACCCAGAGGTGCCCTGAGAAATTCCTCAGCCTCTACATC 3740
Db 3784 ATGATTAATTCAGGCAATCAAGACCCAGAGGTGCCCTGAGAAATTCCTCAGCCTCTACATC 3843
QY 3741 TCTCAGTTTATCATCATGATATAGTCTTGATGGGAAGAGTGGCAGACTTATCGAGGAAAT 3800
Db 3844 TCTCAGTTTATCATCATGATATAGTCTTGATGGGAAGAGTGGCAGACTTATCGAGGAAAT 3903
QY 3801 TCCACTGGAACCTTAATGCTCTCTTTGGCAATGTGATTCATCTGGGATTAACACAAAT 3860
Db 3904 TCCACTGGAACCTTAATGCTCTCTTTGGCAATGTGATTCATCTGGGATTAACACAAAT 3963
QY 3861 ATTTTAAACCTCCCAATTTATGCTCGATACATCCGTTTGCACCCCACTATATAGCAAT 3920
Db 3964 ATTTTAAACCTCCCAATTTATGCTCGATACATCCGTTTGCACCCCACTATATAGCAAT 4023
QY 3921 CGCAGCACTCTTCGATGAGAGTGTAGGGCTGTGATTTAAATAGTTGACGATGCCATTG 3980
Db 4024 CGCAGCACTCTTCGATGAGAGTGTAGGGCTGTGATTTAAATAGTTGACGATGCCATTG 4083
QY 3981 GGAATGAGAGTAAAGCAATATCAGATGACACAGATTACTGCTTCACTTACTTACCAAT 4040
Db 4084 GGAATGAGAGTAAAGCAATATCAGATGACACAGATTACTGCTTCACTTACTTACCAAT 4143
QY 4041 ATGTTGCCACCTGCTCTCTTCAAAAAGCTGCACTTCACTCCAAAGGAGAGTAATGCC 4100
Db 4144 ATGTTGCCACCTGCTCTCTTCAAAAAGCTGCACTTCACTCCAAAGGAGAGTAATGCC 4203
QY 4101 TGGAGCCTCAGGTGAATTAATCCAAAAGAGTGGCTGCAAGTGGACTTCCAGAAGACAATG 4160
Db 4204 TGGAGCCTCAGGTGAATTAATCCAAAAGAGTGGCTGCAAGTGGACTTCCAGAAGACAATG 4263
QY 4161 AAGTACAGAGTAAGTAATCTCAGGAGTAATAATCTGCTTACAGCATGTATGTAAG 4220
Db 4264 AAGTACAGAGTAAGTAATCTCAGGAGTAATAATCTGCTTACAGCATGTATGTAAG 4323
QY 4221 GAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTTTTTCAGAAATGCC 4280
Db 4324 GAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCTTTTTCAGAAATGCC 4383
QY 4281 AAAGTAAAGGTTTTCAGGGAATCAAGACTCCTTCAACCTGTGTGAACCTCTAGAC 4340
Db 4384 AAAGTAAAGGTTTTCAGGGAATCAAGACTCCTTCAACCTGTGTGAACCTCTAGAC 4443
QY 4341 CCACCGTTACTGACTGCTACCTTGAATTCACCCCGAGAGTTGGTTCACACGATTTGCC 4400
Db 4444 CCACCGTTACTGACTGCTACCTTGAATTCACCCCGAGAGTTGGTTCACACGATTTGCC 4503
QY 4401 CTGAGATGAGGTTCTGGGCTGCGAGGCACAGAGACTCTAC 4442
Db 4504 CTGAGATGAGGTTCTGGGCTGCGAGGCACAGAGACTCTAC 4545

RESULT 11
N81545
ID N81545 standard; DNA; 4616 BP.

AC N81545:
DT 04-DEC-1990 (first entry)
DE Human Factor VIII-C analog having exon 4 and 14 deleted.
KW Human Factor VIII-C analog; exon deletion; coagulation disorders;
haemophilia; ss.
PN EP-265778-A.
PD 04-MAY-1988.
PF 14-OCT-1987; 115043.
PR 15-OCT-1986; US-919153.
PA (RORE) Rorer Int Overseas.
PI Sarver N, Drohan W.:
DR MPI: 88-120930/18.
P1 Human Factor VIII-C analogue free of other proteins - produced from
PT recombinant DNA and used in the treatment of coagulation disorders.
PS Claim 4, Page 13; 42pp; English.
CC The genetically engineered analogue can provide a dependable and
readily available therapeutic agent to be used in the treatment of
haemophilia and coagulation disorders in humans or animals.
See also N81543-45.
Sequence 4616 BP; 1333 A; 1038 C; 1012 G; 1232 T;

Query Match 80.2%; Score 3877.2; DB 1; length 4616;
Best Local Similarity 90.4%; Pred. No. 0;
Matches 4366; Conservative 1; Mismatches 9; Indels 454; Gaps 3;

QY 72 ATGCAATAGAGCTCTCCACCTGCTTCTTCTGCGCTTTGGGATTCTGCTTTAGTGCC 131
DB 1 ATGCAATAGAGCTCTCCACCTGCTTCTTCTGCGCTTTGGGATTCTGCTTTAGTGCC 60
QY 132 ACCAGAAGATACCTCTGGTGAGTGAAGTGTATGGAAGTATGCAAGTATGATCTC 191
DB 61 ACCAGAAGATACCTCTGGTGAGTGAAGTGTATGGAAGTATGCAAGTATGATCTC 120
QY 192 GGTGAGCTGCTGTGAGCGCAAGATTCTCTAGAGTGCAGAAATCTTTCCATTCAAC 251
DB 121 GGTGAGCTGCTGTGAGCGCAAGATTCTCTAGAGTGCAGAAATCTTTCCATTCAAC 180
QY 252 ACCTCAGTCGTGTACAAAAAGCTCTGTTGTAGAATTACGGATCACCCTTTTCAACATC 311
DB 181 ACCTCAGTCGTGTACAAAAAGCTCTGTTGTAGAATTACGGATCACCCTTTTCAACATC 240
QY 312 GCTAAGCCAGGCCACCCTGGATGGTCTGCTAGGTCTTACATCCAGGCTGAGTTTAT 371
DB 241 GCTAAGCCAGGCCACCCTGGATGGTCTGCTAGGTCTTACATCCAGGCTGAGTTTAT 300
QY 372 GATACAGTGTCTATTACACTTAAGAATGCTTCCCATCTGTGAGTCTTCACTGCTGT 431
DB 301 GATACAGTGTCTATTACACTTAAGAATGCTTCCCATCTGTGAGTCTTCACTGCTGT 360
QY 432 GGTGATCTCTACTGGAAGCTTCTGAGGAGCTGAATATGATGATCAGACCAAGTCAAGG 491
DB 361 GGTGATCTCTACTGGAAGCTTCTGAGGAGCTGAATATGATGATCAGACCAAGTCAAGG 388
QY 492 GAGAAGAAGATGATAAGTCTTCCCTGGTGAAGCCATATATATGCTGGCAGGCTCTG 551
DB 388 ----- 388
QY 552 AAAGAGATGTGTCATGCGCTCTGACCCACTGCGCTTACCTACTCATATCTTCTCAT 611
DB 388 ----- 388
QY 612 GTGACCTGTGTAAGACTTGAATTCAGGCTCATTTGAGCCCTACTAGTATGTAGAGAA 671
DB 388 ----- 388
QY 672 GGGAGTCTGCCAAGGAAAGACACAGACCTTGACAAATTTATACTACTTTTGTGCTGA 731
DB 388 GGGAGTCTGCCAAGGAAAGACACAGACCTTGACAAATTTATACTACTTTTGTGCTGA 447
QY 732 TTGATGAAGGAAAGTTGGCACTCAGAAACAAAGAACTCTTGATGAGAGATAGGAT 791
DB 448 TTGATGAAGGAAAGTTGGCACTCAGAAACAAAGAACTCTTGATGAGAGATAGGAT 506

QY 792 GCTGCATCTGCTCGGGCTGGCCCTAAATGCAACAGATCATGTTATGTAACAGGCTCT 851
DB 507 GCTGCATCTGCTCGGGCTGGCCCTAAATGCAACAGATCATGTTATGTAACAGGCTCT 566
QY 852 CTGCCAGTCTGATTGGATGCCACAGAAATCAGTCTATGTCATGTGATGGAATGGCC 911
DB 567 CTGCCAGTCTGATTGGATGCCACAGAAATCAGTCTATGTCATGTGATGGAATGGCC 626
QY 912 ACCACTCTGAAGTGCACATATATCTGGAAGTGCACATTTCTGTGAGGAACCAT 971
DB 627 ACCACTCTGAAGTGCACATATATCTGGAAGTGCACATTTCTGTGAGGAACCAT 686
QY 972 CGCCAGGCTCTTGAATCTCGCCAAATTAATCTTACTGCTCAAACTCTTGATG 1031
DB 687 CGCCAGGCTCTTGAATCTCGCCAAATTAATCTTACTGCTCAAACTCTTGATG 746
QY 1032 GACCTTGGACAGTTTCTACTGTTTGTGATATCTTCCCAACCAATGATGCAATGAA 1091
DB 747 GACCTTGGACAGTTTCTACTGTTTGTGATATCTTCCCAACCAATGATGCAATGAA 806
QY 1092 GCTTATGCAAGTAGACAGCTGTCCAGAGGAACCCCACTAGCAATGAAATATGAA 1151
DB 807 GCTTATGCAAGTAGACAGCTGTCCAGAGGAACCCCACTAGCAATGAAATATGAA 866
QY 1152 GAAGCGAAGACTATGATGATGATCTTACTGATTCGAATGATGATGTCAGGTTGAT 1211
DB 867 GAAGCGAAGACTATGATGATGATCTTACTGATTCGAATGATGATGTCAGGTTGAT 926
QY 1212 GATGACACTCTCTCTCTTATCCAAATTCGCTAGTTCGAAGATCCTTAAACT 1271
DB 927 GATGACACTCTCTCTCTTATCCAAATTCGCTAGTTCGAAGATCCTTAAACT 986
QY 1272 TGGTACATTACATTCGCTGTAAGAGAGAGACTGGACTATGCTCCCTAGTCCGCC 1331
DB 987 TGGTACATTACATTCGCTGTAAGAGAGAGACTGGACTATGCTCCCTAGTCCGCC 1046
QY 1332 CCCGATGACAGAAAGTTAAAGTCAATTTGAACAATGGCCCTCAGCGGATGGTAGG 1391
DB 1047 CCCGATGACAGAAAGTTAAAGTCAATTTGAACAATGGCCCTCAGCGGATGGTAGG 1106
QY 1392 AAGTACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTAAAGCTCGTGAAGCT 1451
DB 1107 AAGTACAAAAAGTCCGATTTATGGCATACACAGATGAACCTTAAAGCTCGTGAAGCT 1166
QY 1452 ATTACGATGAATCAGGAATCTTGGACCTTATGGAAGTGGAGACACTG 1511
DB 1167 ATTACGATGAATCAGGAATCTTGGACCTTATGGAAGTGGAGACACTG 1226
QY 1512 TTGATATATTTAAGATCAAGACAGACCAATTAACATCTACCGGATCACT 1571
DB 1227 TTGATATATTTAAGATCAAGACAGACCAATTAACATCTACCGGATCACT 1286
QY 1572 GATGCCGCTCTTGTATTCAGGAGATTACCAAAAGGTGAACAATTTGAAGATTTT 1631
DB 1287 GATGCCGCTCTTGTATTCAGGAGATTACCAAAAGGTGAACAATTTGAAGATTTT 1346
QY 1632 CCAATTCGCGAGAGAAATATTAATATGAAGAGTGAAGTGAAGTGGCCA 1691
DB 1347 CCAATTCGCGAGAGAAATATTAATATGAAGAGTGAAGTGAAGTGGCCA 1406
QY 1692 ACTAAATCAGATCTGGTGGCTGACCCGCTATTACTCTAGTTGTTAATAGAGAGA 1751
DB 1407 ACTAAATCAGATCTGGTGGCTGACCCGCTATTACTCTAGTTGTTAATAGAGAGA 1466
QY 1752 GATCTAGCTTCAGAGACTCATGCGCTCTCTCATCTGCTACAAAGATCTGATGATCAA 1811
DB 1467 GATCTAGCTTCAGAGACTCATGCGCTCTCTCATCTGCTACAAAGATCTGATGATCAA 1526
QY 1812 AGAGGAAACCATATATGTACAGACAGAGAAATGTCATCTGTTTCTGATTTGATGAG 1871
DB 1527 AGAGGAAACCATATATGTACAGACAGAGAAATGTCATCTGTTTCTGATTTGATGAG 1586

QY 1872 AACCGAAGCTGGTACCTCACAGAGATATACAAGCCTTCTCCCAATCCAGCTGGAGTG 1931
|||||
Db 1587 AACCGAAGCTGGTACCTCACAGAGATATACAAGCCTTCTCCCAATCCAGCTGGAGTG 1646
QY 1932 CAGCTTGAGGATCCAGAGTTCACAGCCTCCACATCATGACACAGCATCATGCTATGTT 1991
|||||
Db 1647 CAGCTTGAGGATCCAGAGTTCACAGCCTCCACATCATGACACAGCATCATGCTATGTT 1706
QY 1992 TTTGATAGTTTGCAGTGTGTGAGTTTGTGATGAGGTGGCATACTGGTACATTTCTAAGC 2051
|||||
Db 1707 TTTGATAGTTTGCAGTGTGTGAGTTTGTGATGAGGTGGCATACTGGTACATTTCTAAGC 1766
QY 2052 ATTGAGCAGACACTGACTTCCTTCTGTCTTCTTCTGTGATATACCTTCAACACAAA 2111
|||||
Db 1767 ATTGAGCAGACACTGACTTCCTTCTGTCTTCTTCTGTGATATACCTTCAACACAAA 1826
QY 2112 ATGCTATGAAAGACACACACTACCCCTATTTCCCATTTCTCAGAGAAACTGTCTCATGTGC 2171
|||||
Db 1827 ATGCTATGAAAGACACACACTACCCCTATTTCCCATTTCTCAGAGAAACTGTCTCATGTGC 1886
QY 2172 ATGGAACACCCAGGCTCATGAGATTCTGGGGTGCACAACTCAGACTTTGGAAACAGAGGC 2231
|||||
Db 1887 ATGGAACACCCAGGCTCATGAGATTCTGGGGTGCACAACTCAGACTTTGGAAACAGAGGC 1946
QY 2232 ATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAAAGAACTGCTGATTATACGAGAC 2291
|||||
Db 1947 ATGACCGCCTTACTGAAGGTTTCTAGTTGTGACAAAGAACTGCTGATTATACGAGAC 2006
QY 2292 AGTTATGAAGATATTTACAGCATCTTGTCTGAGTAAACAAATGCCATTGAACCAAGAGC 2351
|||||
Db 2007 AGTTATGAAGATATTTACAGCATCTTGTCTGAGTAAACAAATGCCATTGAACCAAGAGC 2066
QY 2352 TTCTCCAG----- 2361
|||||
Db 2067 TTCTCCAGAAATTCATATTTGATCTCTTGTGGAATAACCACTATGTACTCAGATA 2126
QY 2361 ----- 2361
Db 2127 CCAAAAGAAAGTGGAATCCCAAGAGATCACAGAAAAACAGCTTTTAAGAAAAAG 2186
QY 2361 ----- 2361
Db 2187 GATACCAATTTGTCCCTGACGCTTGTGAAGCAATCATGCAATAGCAGCAATAATGAG 2246
QY 2361 ----- 2361
Db 2247 GGACAAAATTAAGCCGAAATAGAAGTCACTGGGCAAAAGCAAGGTAGGACTGAAGGCTG 2306
QY 2361 -----AACCCACAGTCTTGAACGCCATCAACGGGAAATTAAGTACTACTCTT 2411
|||||
Db 2307 TGCTCTCAAAACCCACACAGCTTGAACGCCATCAACGGGAAATTAAGTACTACTCTT 2366
QY 2412 CAGTCAGATCAAGAGAAATGACTATGATGATACCATATCAGTTGAATGAAGAAGAA 2471
|||||
Db 2367 CAGTCAGATCAAGAGAAATGACTATGATGATACCATATCAGTTGAATGAAGAAGAA 2426
QY 2472 GATTTGACATTTATGATGAGGATGAATAATCAGAGCCCGCAGCTTTCAAAAGAAACA 2531
|||||
Db 2427 GATTTGACATTTATGATGAGGATGAATAATCAGAGCCCGCAGCTTTCAAAAGAAACA 2486
QY 2532 CGACACTATTTATTTGCTGAGTGAGAGGCTCTGGGATTTAGGATGAGTACTCCCA 2591
|||||
Db 2487 CGACACTATTTATTTGCTGAGTGAGAGGCTCTGGGATTTAGGATGAGTACTCCCA 2546
QY 2592 CATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTC 2651
|||||
Db 2547 CATGTTCTAAGAAACAGGGCTCAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGTTGTTTC 2606
QY 2652 CAGGAATTTACTGATGGCTTTACTCAGCCCTTATACGTTGGAGAACTAAATGAACAT 2711
|||||
Db 2607 CAGGAATTTACTGATGGCTTTACTCAGCCCTTATACGTTGGAGAACTAAATGAACAT 2666
QY 2712 TTGGACTCTGGGGCATATATAAGACAGAAAGTTGAAGATAATATCATGTACTTTC 2771

Db 2667 TTGGACTCTCTGGGGCATATATAAGACAGAAAGTTGAAGATAATATCATGTACTTTC 2726
QY 2772 AGAAATCAGGCCTCTCGTCCCTATTTCTCTATTTCTAGCCTTATTTCTATGAGGAAGAT 2831
|||||
Db 2727 AGAAATCAGGCCTCTCGTCCCTATTTCTCTATTTCTAGCCTTATTTCTATGAGGAAGAT 2786
QY 2832 CAGAGGCAAGGAGCAGAACTAGAAAAAATTTGTCAAGCCTAATGAACCAAACTTAC 2891
|||||
Db 2787 CAGAGGCAAGGAGCAGAACTAGAAAAAATTTGTCAAGCCTAATGAACCAAACTTAC 2846
QY 2892 TTTGGAAGTGCACATCATATATGGCACCCACTAAAGATGAGTTTACTGCAAAAGCCTTG 2951
|||||
Db 2847 TTTTGAAGTGCACATCATATATGGCACCCACTAAAGATGAGTTTACTGCAAAAGCCTTG 2906
QY 2952 GCTTATTTCTCTGATGTGACCTGGAAGAAAGATGTGCACCTCAGGCCCTGATGGACCCCTT 3011
|||||
Db 2907 GCTTATTTCTCTGATGTGACCTGGAAGAAAGATGTGCACCTCAGGCCCTGATGGACCCCTT 2966
QY 3012 CTGCTGCGCACACTAACAACACTGAACCCCTGCTCATGAGAGACAAAGTACAGAGAA 3071
|||||
Db 2967 CTGCTGCGCACACTAACAACACTGAACCCCTGCTCATGAGAGACAAAGTACAGAGAA 3026
QY 3072 TTTGCTCTGTTTTCACCATCTTTGATGAGAACCAAAAGCTGTTACTTCACTGAAAATATG 3131
|||||
Db 3027 TTTGCTCTGTTTTCACCATCTTTGATGAGAACCAAAAGCTGTTACTTCACTGAAAATATG 3086
QY 3132 GAAAGAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAAAT 3191
|||||
Db 3087 GAAAGAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAAAT 3146
QY 3192 TATGCTTCCATGCAATATGCTTACATATATGATATACCTACCTGGCTTAATGGCT 3251
|||||
Db 3147 TATGCTTCCATGCAATATGCTTACATATATGATATACCTACCTGGCTTAATGGCT 3206
QY 3252 CAGGATCAAAGGATTCGATGGTATCTGCTCAGCATGGGCAACAAATGAAGAACTCATCT 3311
|||||
Db 3207 CAGGATCAAAGGATTCGATGGTATCTGCTCAGCATGGGCAACAAATGAAGAACTCATCT 3266
QY 3312 ATTCATTTCAAGTGCATGTGTTCACTGTACGAAAAAAGAGAGTATTAATGGCACTG 3371
|||||
Db 3267 ATTCATTTCAAGTGCATGTGTTCACTGTACGAAAAAAGAGAGTATTAATGGCACTG 3326
QY 3372 TACAATCTCTATCCAGGTGTTTGTGAGACAGTGGAATGTTACCATCCAAAGCTGGAAT 3431
|||||
Db 3327 TACAATCTCTATCCAGGTGTTTGTGAGACAGTGGAATGTTACCATCCAAAGCTGGAAT 3386
QY 3432 TGGCGGTGGAATGCTTATTTGGGAGGATCTACATGCTGGATGAGCACACTTTTCTG 3491
|||||
Db 3387 TGGCGGTGGAATGCTTATTTGGGAGGATCTACATGCTGGATGAGCACACTTTTCTG 3446
QY 3492 GTGTACAGCAATTAAGTGTACAGACTCCCTGGGAATGCTTCTGACACATTTAGAGATTT 3551
|||||
Db 3447 GTGTACAGCAATTAAGTGTACAGACTCCCTGGGAATGCTTCTGACACATTTAGAGATTT 3506
QY 3552 CAGATTACGCTTCAGGACAAATATGAGACGTGGGCCCAAAAGCTGGCCAGACTTCATTAT 3611
|||||
Db 3507 CAGATTACGCTTCAGGACAAATATGAGACGTGGGCCCAAAAGCTGGCCAGACTTCATTAT 3566
QY 3612 TCCGATCAATCAATGCTTGAGACACCAAGAGCCCTTTCTTGTGATCAAGGTGATCTG 3671
|||||
Db 3567 TCCGATCAATCAATGCTTGAGACACCAAGAGCCCTTTCTTGTGATCAAGGTGATCTG 3626
QY 3672 TTGGACCAATGATTTATCAAGGACATCAAGACCAGGGTGCCTCAGAAAGTTCTCCAGC 3731
|||||
Db 3627 TTGGACCAATGATTTATCAAGGACATCAAGACCAGGGTGCCTCAGAAAGTTCTCCAGC 3686
QY 3732 CTCTACATCTCTCAGTTTATCATCATGATATAGTCTTGAATGGGAAGAAAGTGGCAGACTTAT 3791
|||||
Db 3687 CTCTACATCTCTCAGTTTATCATCATGATATAGTCTTGAATGGGAAGAAAGTGGCAGACTTAT 3746
QY 3792 CGAGGAATTCACACTGGAACCTTAATGCTCTTCTTGGCAATGTGATTCATCTGGGATA 3851
|||||

Db 3747 CGAGGAATTCACCTGGAACTTAATGGTCTTCTTTGGCAATGTGATTCATCTGGGATA 3806
QY 3852 AAACACAAATATTTTAAACCCCTCCAAATATTGGCTGATACATCCGTTTGACCACTCAT 3911
Db 3807 AAACACAAATATTTTAAACCCCTCCAAATATTGGCTGATACATCCGTTTGACCACTCAT 3866
QY 3912 TATAGCATTCGAGCACTCTTCGATGAGTGTGGCTGTGATTTAAATAGTTCAGC 3971
Db 3867 TATAGCATTCGAGCACTCTTCGATGAGTGTGGCTGTGATTTAAATAGTTCAGC 3926
QY 3972 ATGCCATTGGGAATGGAGATGAAGCAATATCAGATGCACAGATTAATCTGCTCATCTAC 4031
Db 3927 ATGCCATTGGGAATGGAGATGAAGCAATATCAGATGCACAGATTAATCTGCTCATCTAC 3986
QY 4032 TTACCAATATGTTTGGCCACCTGGTCTCTTCAAAAAGCTCGACTTCACCTCCAGGGAGG 4091
Db 3987 TTACCAATATGTTTGGCCACCTGGTCTCTTCAAAAAGCTCGACTTCACCTCCAGGGAGG 4046
QY 4092 AGTAATGCTGGAGACCTCAGGTGAATAATCCAAAAGAGTGGCTGCAAGTGGACTTCCAG 4151
Db 4047 AGTAATGCTGGAGACCTCAGGTGAATAATCCAAAAGAGTGGCTGCAAGTGGACTTCCAG 4106
QY 4152 AAGACAATGAAGTCCACAGAGATACTACTCAGGAGTAATAATCTCTGCTTACCAAGCATG 4211
Db 4107 AAGACAATGAAGTCCACAGAGATACTACTCAGGAGTAATAATCTCTGCTTACCAAGCATG 4166
QY 4212 TATGTGAAGAGTCTCTCATCTCCAGCAGATCAAGATGGCCATCAGTGGACTCTCTTTT 4271
Db 4167 TATGTGAAGAGTCTCTCATCTCCAGCAGATCAAGATGGCCATCAGTGGACTCTCTTTT 4226
QY 4272 CAGAATGGCAAGTAAGGTTTTCAGGGAATCAAGACTCTTCACACCTGTGTGAAC 4331
Db 4227 CAGAATGGCAAGTAAGGTTTTCAGGGAATCAAGACTCTTCACACCTGTGTGAAC 4286
QY 4332 TCTCTAGACCCACCGTTACTGACTGCTACTCTCGAATTCACCCCGAGAGTGGTGAC 4391
Db 4287 TCTCTAGACCCACCGTTACTGACTGCTACTCTCGAATTCACCCCGAGAGTGGTGAC 4346
QY 4392 CAGATTGCCCTGAGGATGAGGTTCTGGGCTGGGAGGACACAGCACTCTACTGAGGTGG 4451
Db 4347 CAGATTGCCCTGAGGATGAGGTTCTGGGCTGGGAGGACACAGCACTCTACTGAGGTGG 4406
QY 4452 CCACTGACAGCACTGCCACTGCCCTCCCTCCCTCCTCAGCTCCAGGGCAGTGTCCCT 4511
Db 4407 CCACTGACAGCACTGCCACTGCCCTCCCTCCCTCCTCAGCTCCAGGGCAGTGTCCCT 4466
QY 4512 CCCTGGCTTGCCTTCTACCTTTGTGCTAAATCCCTAGCAGACACTGCCTTGAAGCCTCCTG 4571
Db 4467 CCCTGGCTTGCCTTCTACCTTTGTGCTAAATCCCTAGCAGACACTGCCTTGAAGCCTCCTG 4526
QY 4572 AATTAACTATCATCAGTCCCTGATTTCTTTGGTGGGGGCGCAGAGGGTGCATCCAAATT 4631
Db 4527 AATTAACTATCATCAGTCCCTGATTTCTTTGGTGGGGGCGCAGAGGGTGCATCCAAATT 4586
QY 4632 AACTTAACCTTACCTATTTCTGACGCTG 4661
Db 4587 AACTTAACCTTACCTATTTCTGACGCTG 4616

RESULT 12
ID T69811 standard; DNA; 5035 BP.
AC T69811;
DT 10-AUG-1997 (first entry)
DE Factor VIII-dB695-HCII DNA.
KW Factor VIII-dB695-HCII; heparin cofactor II; blood coagulation;
KW blood clotting; procoagulant; anticoagulant; antithrombotic;
KW haemophilia; gene therapy; ss.
OS Homo sapiens.
FH key Location/Qualifiers
FT cds 35..5020
FT /tag= a
FT /product= Factor VIII-dB695-HCII

FT misc_rna 2225..2314
FT /tag= b
FT /product= heparin cofactor II region (aa51-81)
PN WO9718315-A1.
PD 22-MAY-1997.
PF 13-NOV-1996; E04977.
PR 13-NOV-1995; US-558107.
PA (IMMO) IMMUNO AG.
PI Voorberg JJ;
DR WPI: 97-289291/26.
DR P-PSDB; W18670.
PT Hybrid Factor VIII with modified activity, comprises region from
PT donor anticoagulant or antithrombotic protein - useful for treatment
PT of coagulation disorders
PS Claim 16; Page 52-60; 96pp; English.
CC A DNA molecule (T69811) codes for Factor VIII-dB695-HCII (W18670),
CC a hybrid protein in which amino acids 712-736 of Factor-dB695
CC (Factor VIII del868-1562) B-domain are replaced by amino acids
CC 51-80 from the acidic region (and potential thrombin-binding site)
CC of human heparin cofactor II (HCII). It was obtd. by PCR
CC amplification (see also T69812-13) of the HCII acidic region from
CC total liver cDNA, fusion to sequences encoding Factor VIII
CC aa706-711 and aa737-743, and incorporation of the construct into
CC plasmid pCLB-dB695. The hybrid protein, which can be expressed
CC using gene therapy techniques, has increased procoagulant activity
CC owing to the HCII acidic region, and can be used to treat blood
CC coagulation disorders such as haemophilia A.
SQ Sequence 5035 BP; 1484 A; 1127 C; 1110 G; 1314 T;

Query Match 76.5%; Score 3696.4; DB 1; Length 5035;
Best Local Similarity 86.7%; Pred. No. 0;
Matches 4362; Conservative 0; Mismatches 56; Indels 612; Gaps 2;

QY 43 CTCAGTTGACATTTGTAGCAATAAGTACGCAATAGAGCTCTCCACCTGCTCTTTC 102
Db 6 CTCAGTTGACATTTGTAGCAATAAGTACGCAATAGAGCTCTCCACCTGCTCTTTC 65
QY 103 TGTGCTTTTGGCATTCGCTTTAGTGGCCACCGAAGATACTACCTGGGTGACGTGAAC 162
Db 66 TGTGCTTTTGGCATTCGCTTTAGTGGCCACCGAAGATACTACCTGGGTGACGTGAAC 125
QY 163 TGTGATGGGACTATATGCAAGTATCGGTGAGCTGCTGTGAGCGAAGATTTCCTC 222
Db 126 TGTGATGGGACTATATGCAAGTATCGGTGAGCTGCTGTGAGCGAAGATTTCCTC 185
QY 223 CTAGAGTGGCAAAATCTTTTCCATTCAACACCTCAGTGTGTACAAAAGACTGTGTTG 282
Db 186 CTAGAGTGGCAAAATCTTTTCCATTCAACACCTCAGTGTGTACAAAAGACTGTGTTG 245
QY 283 TAGAATTCACGGATCACCTTTTCAACATCGCTAAGCCAAAGGCCACCTGGATGGGTGCG 342
Db 246 TAGAATTCACGGATCACCTTTTCAACATCGCTAAGCCAAAGGCCACCTGGATGGGTGCG 305
QY 343 TAGTCTTACCATCCAGGCTGAGGTTATGATACAGTGTGATTAAGTCTTCCCTGGTG 402
Db 306 TAGTCTTACCATCCAGGCTGAGGTTATGATACAGTGTGATTAAGTCTTCCCTGGTG 365
QY 403 CTCCCATCCTGTCACTCTCATGCTGTGTTGTTATCTACTGGAAGCTTCTGAGGGAG 462
Db 366 CTCCCATCCTGTCACTCTCATGCTGTGTTGTTATCTACTGGAAGCTTCTGAGGGAG 425
QY 463 CTGAATATGATGATCAAGCCAGTCAAAAGGAGAAAGATGATAAGTCTTCCCTGGTG 522
Db 426 CTGAATATGATGATCAAGCCAGTCAAAAGGAGAAAGATGATAAGTCTTCCCTGGTG 485
QY 523 GAAGCCATACATATGTCTGGCAGGTCTGAAAAGAGATGTCCAAATGGCTCTGACCCAC 582
Db 486 GAAGCCATACATATGTCTGGCAGGTCTGAAAAGAGATGTCCAAATGGCTCTGACCCAC 545
QY 583 TGTGCTTACCTACTCATATCTTCTCATGTGAGACCTGGTAAAGACTTGAATTCAGGCC 642
Db 546 TGTGCTTACCTACTCATATCTTCTCATGTGAGACCTGGTAAAGACTTGAATTCAGGCC 605

QY	643	TCATGGAGCCCTACTAGTATGTAGAGAGGGAGTCTGGCCAAAGGAAAAGACACAGACCT	702
Db	606	TCATTGGAGCCCTACTAGTATGTAGAGAGGGAGTCTGGCCAAAGGAAAAGACACAGACCT	665
QY	703	TGCACAATTATACTACTTTTGTGCTGATTGTGATGAAGGGAAAAGTTGGCACTCAGAAA	762
Db	666	TGCACAATTATACTACTTTTGTGCTGATTGTGATGAAGGGAAAAGTTGGCACTCAGAAA	725
QY	763	CAAGAAGCTCCTTGATGACGAGTAGGGATGCTGCATCTGCTCGGGCCCTGAAATGC	822
Db	726	CAAGAAGCTCCTTGATGACGAGTAGGGATGCTGCATCTGCTCGGGCCCTGAAATGC	785
QY	823	ACACAGTCAATGGTTATGTAAACAGGTCCTCGCCAGGTCGTGATTGGATGCCACAGAAAT	882
Db	786	ACACAGTCAATGGTTATGTAAACAGGTCCTCGCCAGGTCGTGATTGGATGCCACAGAAAT	845
QY	883	CAGTCTATTGGCATGTGATTGGAATGGGCAACCACCTCCTGAAAGTGCACATATTCTCG	942
Db	846	CAGTCTATTGGCATGTGATTGGAATGGGCAACCACCTCCTGAAAGTGCACATATTCTCG	905
QY	943	AAGGTACACACATTTCTTGTGAGGAACCATCGCCAGGCGCTTGGAATCTCGCCAAATA	100
Db	906	AAGGTACACACATTTCTTGTGAGGAACCATCGCCAGGCGCTTGGAATCTCGCCAAATA	965
QY	1003	CTTTCCTTACTGCTCTCAACACACTCTGATGAGACCTTGACAGTTTCTACTGTTTGTCTATA	106
Db	966	CTTTCCTTACTGCTCTCAACACACTCTGATGAGACCTTGACAGTTTCTACTGTTTGTCTATA	102
QY	1063	TCTCTTCCCACCAACATGATGGCATGGAAGCTTATGTCAAAGTAGACACCTGTCCAGAG	112
Db	1026	TCTCTTCCCACCAACATGATGGCATGGAAGCTTATGTCAAAGTAGACACCTGTCCAGAG	108
QY	1123	AACCCCACTACGAATGAAAAAATAATGAGAAGCGGAAGACTATGATGATCTTACTG	118
Db	1086	AACCCCACTACGAATGAAAAAATAATGAGAAGCGGAAGACTATGATGATCTTACTG	114
QY	1183	ATTCTGAATGATGTGCTCAGGTTTGATGATGACAACTCTCCTTATCCAAATTC	124
Db	1146	ATTCTGAATGATGTGCTCAGGTTTGATGATGACAACTCTCCTTATCCAAATTC	120
QY	1243	GCTCAGTTGCCAAGAAGCATCTAAAACTGGGTACATTACATTTGCTGCTGAAGAGAG	130
Db	1206	GCTCAGTTGCCAAGAAGCATCTAAAACTGGGTACATTACATTTGCTGCTGAAGAGAG	126
QY	1303	ACTGGGACTATGCTCCCTTAGTCCCTCGCCCCGATGACAGAAGTTATATAAGTCAATATT	136
Db	1266	ACTGGGACTATGCTCCCTTAGTCCCTCGCCCCGATGACAGAAGTTATATAAGTCAATATT	132
QY	1363	TGAACAATGGCCCTCAGCGGATTTGTAAGAAAGTACAAAAAAAGTCCGATTATGGCATACA	142
Db	1326	TGAACAATGGCCCTCAGCGGATTTGTAAGAAAGTACAAAAAAAGTCCGATTATGGCATACA	138
QY	1423	CAGATGAAACCTTTAAGACTCGTGAAGCTATTGAGCATGAATCAGGAATCTTGGACCTT	148
Db	1386	CAGATGAAACCTTTAAGACTCGTGAAGCTATTGAGCATGAATCAGGAATCTTGGACCTT	144
QY	1483	TACTTTATGGGGAAGTTGGAGACACACACTGTTGATTATATTTAAGAATCAAGCAAGCAGAC	154
Db	1446	TACTTTATGGGGAAGTTGGAGACACACACTGTTGATTATATTTAAGAATCAAGCAAGCAGAC	150
QY	1543	CATATAACATCTACCCCTCACGGAATCACTGATGTCGTCCTTGTATCAAGAGAGATTAC	160
Db	1506	CATATAACATCTACCCCTCACGGAATCACTGATGTCGTCCTTGTATCAAGAGAGATTAC	156
QY	1603	CAAAAGGTGTAAAAACATTTGAAGGATTTTCCAATTCTGCCAGAGAGAAAATTCCAATATA	166
Db	1566	CAAAAGGTGTAAAAACATTTGAAGGATTTTCCAATTCTGCCAGAGAGAAAATTCCAATATA	162
QY	1663	AATGACAGTGAAGTGTAGAAAGATGGGCCAACTAAATCAGATCCTCGGTGCTGACCCGCT	172
Db	1626	AATGACAGTGAAGTGTAGAAAGATGGGCCAACTAAATCAGATCCTCGGTGCTGACCCGCT	168

QY	1723	ATTACTCTAGTTTCGTTTAATATGAGAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCC	1782
Db	1686	ATTACTCTAGTTTCGTTTAATATGAGAGAGATCTAGCTTCAGGACTCATTTGGCCCTCTCC	1745
QY	1783	TCATCTGCTACAAAGAAATCTGTAGATCCAAAGAGAAAACAGATTAATGTCCAGACAAGAGA	1842
Db	1746	TCATCTGCTACAAAGAAATCTGTAGATCCAAAGAGAAAACAGATTAATGTCCAGACAAGAGA	1805
QY	1843	ATGTCACTCTGTTTCTCTGTATTGATGAGAACCGAAGCTGTACTCTCACAGAAATATAC	1902
Db	1806	ATGTCACTCTGTTTCTCTGTATTGATGAGAACCGAAGCTGTACTCTCACAGAAATATAC	1865
QY	1903	AACGCTTCTCCCAATCCAGCTGGAGTGCAGCTTGAGATCCAGAGTTCAGCCTCCA	1962
Db	1866	AACGCTTCTCCCAATCCAGCTGGAGTGCAGCTTGAGATCCAGAGTTCAGCCTCCA	1925
QY	1963	ACATCATGCGACAGCATCAATGGCTAATGTTTGTAGTATGTTTGCAGTTGTCAAGCTTGGC	2022
Db	1926	ACATCATGCGACAGCATCAATGGCTAATGTTTGTAGTATGTTTGCAGTTGTCAAGCTTGGC	1985
QY	2023	ATGAGGTGGCATACTGCTGTACATTCATAAGCATGGAGCACAGACTGACTTCCTTCTGTCT	2082
Db	1986	ATGAGGTGGCATACTGCTGTACATTCATAAGCATGGAGCACAGACTGACTTCCTTCTGTCT	2045
QY	2083	TCTTCTCTGGATATACCTTCMAACCAAAATGGTCTATGAAGACACACTCACCTTATTCC	2142
Db	2046	TCTTCTCTGGATATACCTTCMAACCAAAATGGTCTATGAAGACACACTCACCTTATTCC	2105
QY	2143	CATTCTCAGGAGAAACGTCTTCATGTCTGATGGAAAAACCAGCTTATGATTTCTGGGT	2202
Db	2106	CATTCTCAGGAGAAACGTCTTCATGTCTGATGGAAAAACCAGCTTATGATTTCTGGGT	2165
QY	2203	GCCACACTCAGACTTTCGGAACAGAGGCGATGACCGCCTTACTGAAGGTTCTAGTTGT-	2262
Db	2166	GCCACACTCAGACTTTCGGAACAGAGGCGATGACCGCCTTACTGAAGGTTCTAGTTGT-	2225
QY	2262	-----GACAAGAACACTGCTGATTATTACGAGGACAGTTATGAGATATTT	2307
Db	2226	TTCCAGAGGGGGAGAGGACGACGACTATCTGACCTGGAGAGATATTACGTGAAGACG	2285
QY	2308	CAGCATACTTGTCTGAGTAAAAACAATGCCATTGAAACCAAGAACTTCTCCAG-----	2361
Db	2286	ACGACTACATCGACATCTGCGACAGTCTGATTGAAACCAAGAACTTCTCCCAATTCAA	2345
QY	2361	-----	2361
Db	2346	GACACCCTAGCACTAGGCAAAAGCAATTTAATGCCACCACAATTCAGAAAAATGACATAG	2405
QY	2361	-----	2361
Db	2406	AGAGACTGACCCTTGGTTTGCACACAGAACACACTATGCTTAATAATACMAAAATGTCTCT	2465
QY	2361	-----	2361
Db	2466	CTAGTGAATTGTTGATGCTCTTGGACAGAGTCTTACTCCACATGGGCTATCTTATCTG	2525
QY	2361	-----	2361
Db	2526	ATCTCCAAGAAAGCAAAATATGAGACTTTTCTGTATGATCCATCACTGAGCAATAGACA	2585
QY	2361	-----	2361
Db	2586	GTAATACAGCCCTGTCTGAATGACACACTTCAGGGCCACAGCTCCATCACAGTGGGACA	2645
QY	2361	-----	2361
Db	2646	TGGTATTTACCCCTGAGTCCAGGCTCCAAATTAAGATTAAATGAGAAACTGGGCACTG	2705
QY	2361	-----	2361
Db	2706	CAGATCCTCTTGCTTGGGATTAACCACTATGTTACTCAGATACCAAAAGAGAGTGAAT	2765
QY	2361	-----	2361

Db 2766 CCCAAGAGAAGTCACCAGAAAAAACAGCTTTTAAAGAAAAAGATACCAATTTTGTCCCTGA 2825
QY 2361 ----- 2361
Db 2826 ACGCTGTGAAGCAATCATGCAATAGCAGCAATAAATGAGGGACAAAATAAGCCCGAAA 2885
QY 2361 -----AACCCACCAG 2370
Db 2886 TAGAAGTCACCTGGCAAGCAAGTAGAGCTGAAAGGCTGTGCTCTCAAAACCACAG 2945
QY 2371 TCTTGAACGCCATCAACGGGAATACTCGTACTACTCTTCAGTCAGATCAAGAGAAA 2430
Db 2946 TCTTGAACGCCATCAACGGGAATACTCGTACTACTCTTCAGTCAGATCAAGAGAAA 3005
QY 2431 TTGACTATGATATACCATATCAGTTTGAATGAAGAAGAGATTTTGACATTATGATG 2490
Db 3006 TTGACTATGATATACCATATCAGTTTGAATGAAGAAGAGATTTTGACATTATGATG 3065
QY 2491 AGGATGAAATCAGAGCCCCCGCAGCTTTCAAAGAAAACACGACACTATTTTATGCTG 2550
Db 3066 AGGATGAAATCAGAGCCCCCGCAGCTTTCAAAGAAAACACGACACTATTTTATGCTG 3125
QY 2551 CAGTGGAGAGGCTCTGGGATTATGGGATGAGTAGCTCCCCACATGTTCTAAGAAAACAGG 2610
Db 3126 CAGTGGAGAGGCTCTGGGATTATGGGATGAGTAGCTCCCCACATGTTCTAAGAAAACAGG 3185
QY 2611 CTGAGATGGCAGTGTCCCTCAGTTTCAAGAAAGTGTTCAGAGAAATTTACTGATGCT 2670
Db 3186 CTGAGATGGCAGTGTCCCTCAGTTTCAAGAAAGTGTTCAGAGAAATTTACTGATGCT 3245
QY 2671 CCTTACTCAGCCCTTATACCGTGAAGAACTAAATGAACATTTGGAGCTCCTGGGCCAT 2730
Db 3246 CCTTACTCAGCCCTTATACCGTGAAGAACTAAATGAACATTTGGAGCTCCTGGGCCAT 3305
QY 2731 ATATAGAGCAGAAGTTGAAGATATATCATGTAATTTTCAGAAATCAGGCCCTCTGTC 2790
Db 3306 ATATAGAGCAGAAGTTGAAGATATATCATGTAATTTTCAGAAATCAGGCCCTCTGTC 3365
QY 2791 CCTATTCCTTCTATTCAGCCTTATTTCTTATGAGAGAGATCAGAGGCAAGGACGAAC 2850
Db 3366 CCTATTCCTTCTATTCAGCCTTATTTCTTATGAGAGAGATCAGAGGCAAGGACGAAC 3425
QY 2851 CTAGAAAAAATTTGTCAAGCCTAATGAACCAAACTTACTTTTGGAAAGTGCAACATC 2910
Db 3426 CTAGAAAAAATTTGTCAAGCCTAATGAACCAAACTTACTTTTGGAAAGTGCAACATC 3485
QY 2911 ATATGGCACCCACTAAAGATGAGTTTGACTGCAAAAGCCTGGGCTTATTTCTGTGTTG 2970
Db 3486 ATATGGCACCCACTAAAGATGAGTTTGACTGCAAAAGCCTGGGCTTATTTCTGTGTTG 3545
QY 2971 ACCTGAAAAAGATGTGCACCTCAGGCTGATTGGACCCCTTCTGTCTGCCACACTAACA 3030
Db 3546 ACCTGAAAAAGATGTGCACCTCAGGCTGATTGGACCCCTTCTGTCTGCCACACTAACA 3605
QY 3031 CACTGAACCCCTGCTCATGGAGACAAGTGACAGTACAGGAATTTGCTCTGTTTTCACCA 3090
Db 3606 CACTGAACCCCTGCTCATGGAGACAAGTGACAGTACAGGAATTTGCTCTGTTTTCACCA 3665
QY 3091 TCTTTGATGAGACCAAAAGCTGTGTTACTTCACTGAAAAATATGAAAAGAACTGCAAGGCTC 3150
Db 3666 TCTTTGATGAGACCAAAAGCTGTGTTACTTCACTGAAAAATATGAAAAGAACTGCAAGGCTC 3725
QY 3151 CCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAAATTATGCTTCCATGCAATCA 3210
Db 3726 CCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAGAAATTATGCTTCCATGCAATCA 3785
QY 3211 ATGGCTACATATGATATACACTACCTGGCTTAGTAATGGCTCAGAGTCAAAAGATTGAT 3270
Db 3786 ATGGCTACATATGATATACACTACCTGGCTTAGTAATGGCTCAGAGTCAAAAGATTGAT 3845
QY 3271 GGTAATCTGCTCAGCATGGGACAGCAATGAAAACATTCATTTCAATTTCAAGTGCATG 3330
|||||

Db 3846 GGTAATCTGCTCAGCATGGGACAGCAATGAAAACATTCATTTCAATTTTCAAGTGCATG 3905
QY 3331 TGTTCACCTGTACGAAAAAAGAGAGATATAAATGGCACCTGTACAACTCTATCCAGGTG 3390
Db 3906 TGTTCACCTGTACGAAAAAAGAGAGATATAAATGGCACCTGTACAACTCTATCCAGGTG 3965
QY 3391 TTTTGGAGACAGTGAATGTATCCATCCAAAGCTGGAATTTGGGGGTGAATGCCCTTA 3450
Db 3966 TTTTGGAGACAGTGAATGTATCCATCCAAAGCTGGAATTTGGGGGTGAATGCCCTTA 4025
QY 3451 TTGGCGAGCATCTACATGCTGGGATGAGACACACTTTTCTGTGTGTACAGCAATAGTGTG 3510
Db 4026 TTGGCGAGCATCTACATGCTGGGATGAGACACACTTTTCTGTGTGTACAGCAATAGTGTG 4085
QY 3511 AGACTCCCTGGGAATGGCTTCTGGACACATTAAGATTTTCAGATTACAGCTTCAGAG 3570
Db 4086 AGACTCCCTGGGAATGGCTTCTGGACACATTAAGATTTTCAGATTACAGCTTCAGAG 4145
QY 3571 AATATGACAGTGGGCCCCCAAGCTGGCCAGACTTCATTTATCCGGATCAATCAATGCC 3630
Db 4146 AATATGACAGTGGGCCCCCAAGCTGGCCAGACTTCATTTATCCGGATCAATCAATGCC 4205
QY 3631 GGAACACCAAGAGGCCCTTTTCTTGATCAAGGTGGATCTGTGGCACCACATGATTATTC 3690
Db 4206 GGAACACCAAGAGGCCCTTTTCTTGATCAAGGTGGATCTGTGGCACCACATGATTATTC 4265
QY 3691 ACGGCATCAAGACCCAGGGTGGCCCTCAGAGTTCCTCCAGCTTACATCTCTCAGTTTA 3750
Db 4266 ACGGCATCAAGACCCAGGGTGGCCCTCAGAGTTCCTCCAGCTTACATCTCTCAGTTTA 4325
QY 3751 TCATCATGTATAGTCTTGATGGGAAGAAGTGGCAGACTTATCGAGGAATTCACCTGAA 3810
Db 4326 TCATCATGTATAGTCTTGATGGGAAGAAGTGGCAGACTTATCGAGGAATTCACCTGAA 4385
QY 3811 CCTTAATGCTCTCTTTGGCAATGTGATTCATCTGGGATATAACACAAATATTTTAAAC 3870
Db 4386 CCTTAATGCTCTCTTTGGCAATGTGATTCATCTGGGATATAACACAAATATTTTAAAC 4445
QY 3871 CTCCAATTATGCTCGATACATCCGTTTGCACCCACTCATTTATAGCATTCGACAGCTC 3930
Db 4446 CTCCAATTATGCTCGATACATCCGTTTGCACCCACTCATTTATAGCATTCGACAGCTC 4505
QY 3931 TTGCGATGAGATTGATGGGCTGTGATTTAAATAGTTGCAGCATGCCATTGGGAATGGAGA 3990
Db 4506 TTGCGATGAGATTGATGGGCTGTGATTTAAATAGTTGCAGCATGCCATTGGGAATGGAGA 4565
QY 3991 GTAAAGCAATATCAGATGCACAGATTACTGCTTCATCTTACCAATATGTTGGCA 4050
Db 4566 GTAAAGCAATATCAGATGCACAGATTACTGCTTCATCTTACCAATATGTTGGCA 4625
QY 4051 CCTGCTCTCTTCAAAAGCTCGACTTCACCTCCAAAGGAGAGATATGCCGTGAGACCTC 4110
Db 4626 CCTGCTCTCTTCAAAAGCTCGACTTCACCTCCAAAGGAGAGATATGCCGTGAGACCTC 4685
QY 4111 AGGTGAATATCCAAAAAGAGTGGCTGCAAGTGACTTCCGAAGACATGAAAAGTCACAG 4170
Db 4686 AGGTGAATATCCAAAAAGAGTGGCTGCAAGTGACTTCCGAAGACATGAAAAGTCACAG 4745
QY 4171 GAGTAATCTACTCAGGAGATATAATCTCTGCTTACCAGCATGTATGTGAAGAGTCTCTCA 4230
Db 4746 GAGTAATCTACTCAGGAGATATAATCTCTGCTTACCAGCATGTATGTGAAGAGTCTCTCA 4805
QY 4231 TCTCCAGCAGTCAAGATGGCCATCAGTGAATCTCTTTTTCAGAAATGGCAAAAGTAAAG 4290
Db 4806 TCTCCAGCAGTCAAGATGGCCATCAGTGAATCTCTTTTTCAGAAATGGCAAAAGTAAAG 4865
QY 4291 TTTTTCAGGGAATCAAGACTCTCTTACACCTGTGGTGAACCTCTCTAGACCACCGTTAC 4350
Db 4866 TTTTTCAGGGAATCAAGACTCTCTTACACCTGTGGTGAACCTCTCTAGACCACCGTTAC 4925
QY 4351 TGACTCGCTACCTTGGAATTCACCCCCAGAGTTGGGTGCACCAGATTGCCCTGAGGATGG 4410
Db 4926 TGACTCGCTACCTTGGAATTCACCCCCAGAGTTGGGTGCACCAGATTGCCCTGAGGATGG 4985
|||||

QY	1569	ACTGATGTCGCTCTTTGTATTC	AAGAGATTACCAAAAGGCTGTA	AAACATTTGAAGAT	1628	
Db	1500	ACTGATGTCAGCGCTTTG	CACCCACGGAGACTTCT	AAAAAGGTGGAAACATTTGAAAGAC	1559	
QY	1629	TTTCCAAATCTG	CCAGAGAAATATTC	CAATATATAATGACAGTGA	CTGTAGAAATGGG	1688
Db	1560	ATGCCAATCTG	CCAGAGAGAGACTTT	CAAGTATAATGACAGTGA	CTGTGGAAGATGGG	1619
QY	1689	CCAACATAATCAGAT	CTCGGTGCTGACCCGCTAT	TACTAGTTCTGTAATATGAG	1748	
Db	1620	CCAACCAAGTCCGAT	CTCGGTGCTGACCCGCTAT	CTACTGAGCTCCATTAATCTAGAG	1679	
QY	1749	AGAGATCTAGCTT	CAAGAGACTCAT	TGGCCCTCTCTCATCTG	TACAAAGAAATCTGTAGAT	1808
Db	1680	AAAGATCTGCTT	CGGGAACTCAT	TGGCCCTCTCTCATCTG	TACAAAGAAATCTGTAGAC	1739
QY	1809	CAAAAGAGAAAC	CAGATTAATGT	CAACAAGAGAAATGT	CAATCTTTCTGTATTTGAT	1868
Db	1740	CAAAAGAGAAAC	CAGATTAATGT	CAACAAGAGAAACGT	CAATCTTTCTGTATTTGAT	1799
QY	1869	GAGAACCGAAGCT	GTGTAACCTC	ACAGAGAAATAT	ACAAGCTTTCTCCCAATCCAGCTGA	1928
Db	1800	GAGAAATCAAAAGCT	GTGTAACCTC	ACAGAGAAATAT	ACAAGCTTTCTCCCAATCCGATGA	1859
QY	1929	GTGCACTTGAGGAT	CCAGATCCAGATTTCCAAAGCT	CTCCACATCATGACAGCATCAATGGCTAT	1988	
Db	1860	TTACAGCCCCAGGAT	CCAGATCCAGATTTCCAAAGCT	CTCCACATCATGACAGCATCAATGGCTAT	1919	
QY	1989	GTTTGTGATAGTT	GTGCAAGTTGTG	TGATGAGGTG	GCATACGTGTAATTTCTA	2048
Db	1920	GTTTGTGATAGTT	GTGCAAGTTGTG	TGATGAGGTG	GCATACGTGTAATTTCTA	1979
QY	2049	AGCATTTGGAGC	ACACAGACTG	ACTTCCTTTCTGCTCTTCTG	ATATACCTTCAACAC	2108
Db	1980	AGTGTGGAGC	ACACAGCGGACTT	CTCTCCGTTCTTCTCTG	GTCTACACCTTCAAAACAC	2039
QY	2109	AAATGCTCTAT	GAAAGACACACT	CAACCTAT	TCCCATTTCTCAGAGAAACGTCTTCATG	2168
Db	2040	AAATGCTCTAT	GAAAGACACACT	CAACCTAT	TCCCATTTCTCAGAGAAACGTCTTCATG	2099
QY	2169	TGCGATGAAAA	CCAGGCTCTAT	GGATTTCTG	GGGTGCACAACATCAGACTTTGGAACAGA	2228
Db	2100	TCAATGGA	AAAAACCCAGGCTCT	TGGGCTCTAGG	GTCCACAACATCAGACTTCCGGAACAGA	2159
QY	2229	GGCATGACCG	CCCTTACTG	AAAGGTTTCTAGTT	GTGACAAGAAACACTGGTATATTACGAG	2288
	2160	GGGATGACAG	CCCTTACTG	AAAGGTTTCTAGTT	GTGACAAGAAACACTGGTATATTACGAG	2219
QY	2289	GACAGTTATGA	AGATATTT	TCAGACATCTT	GTGAGTAAAAACAATGCCATTTGAACCAAGA	2348
Db	2220	AACACTTATGA	AGATATTT	TCAGACATCTT	GTGAGTAAAAACAATGCCATTTGAACCAAGA	2279
QY	2349	AGCTTCTCC	CAGAACCC	CACAGTCTT	GAAACGCCATCAACGGGAAATTAATCTGTACTACT	2408
Db	2280	-----	-----	-----	-----GACATTAAGCCTTCTCTACT	2297
QY	2409	CTTCAGTCAGAT	ATAAGAGAAATTT	GAATGATATAC	CAATATCAGTTGAATGAAGAAG	2468
Db	2298	TTTTCAGCCGG	AGAGAACAAATTT	GACTATGATATATCTT	CTCAACTGAAACGAAGGGA	2357
QY	2469	GAGATTTTGAC	ATTTATGATGAG	GATGAAAAATCAGAGCC	CCGACGTTTCAAAAGAAA	2528
Db	2358	GAGATTTTGAC	ATTTATGATGAG	GATGAAAAATCAGAGCC	CCGACGTTTCAAAAGAAA	2417
QY	2529	ACACGACACT	ATTTATTTATGCT	GACGTGAGAGGCT	CTGGGATTAATGAGATGAGTAGTCTCC	2588
Db	2418	ACCCGACACT	ATTTATTTATGCT	GACGTGAGAGGCT	CTGGGATTAATGAGATGAGTAGTCTCC	2477
QY	2589	CCACATGTTCT	TAAGAAACAGGCGCT	CAGAGTGGCA	GTGCCCTCAGTTCAAGAAAGTTGTT	2648
Db	2478	CCCCGGGCG	CTAAGAAACAGGCGCT	CAGAGTGGCA	GTGCCCTCAGTTCAAGAAAGTTGTT	2537
QY	2649	TTCCAGGAATTA	CTGATGGCTCTT	ACTTACTAGCCCTT	ATAACCGTGAAGAACTAATGAA	2708

Db	2538	TTCCGGGAATTGCTGAGAGGCTCCCTCAGCAGCCGCTGTACCGGGGGAACCAAAA	2597
QY	2709	CATTGGACTCCTGGGCCATATATAAGACAGAGAGTTGAAGATATATCATGGTAACT	2768
Db	2598	CACCTGGGCTCTTGGGACCCTATCATCAGAGCGGAGGTGAAGACAACATCATGTGTAAT	2657
QY	2769	TTCAGAAATCAGGCCCTCTGCTCCCTATTTCTTCTATTTCTAGCCCTATTTCTATAGGAA	2828
Db	2658	TTCAAAAACCAAGCGCTCTCGCTCCCTATTTCTTCTACTCGAGCCCTATTTCTATCCGGAT	2717
QY	2829	GATCAGAGGCAAGGAGCAGAAACCTAGAAAAAACTTTGTCAAGCCCTAATGAACCAAACT	2888
Db	2718	GATCAGAGGCAAGGGGCGAGAAACCTCGACACAACTTCGTCCAGCCCAATGAACCAAGAACT	2777
QY	2889	TACTTTGGAAAAGTGCACATCATATAGGCACCCACTAAAGATGATTTGACTGCAAAAGCC	2948
Db	2778	TACTTTGGAAAAGTGCACATCATATAGGCACCCACTAAAGATGATTTGACTGCAAAAGCC	2837
QY	2949	TGGGCTTATTTCTCTGATGTTGACCTGGAAAAAGATGTGCACTCAGGCCCTGATTTGACCC	3008
Db	2838	TGGGCTTATTTCTCTGATGTTGACCTGGAAAAAGATGTGCACTCAGGCCCTGATTTGACCC	2897
QY	3009	CTTCTGCTCTGCCACACTAACACACTGAAACCCCTGCTCATGGGAGACAAGTGACAGTACAG	3068
Db	2898	CTTCTGATCTGCCCGCGCCCAACACCCCTGAAACGCTGCTCATCGGTAAACAAGTGACCGTCAA	2957
QY	3069	GAATTTGCTCTGTTTTCACCATCTTTGATGAGACCAAAAGCTGCTACTTCACTGAAAAAT	3128
Db	2958	GAATTTGCTCTGTTTTCACCATCTTTGATGAGACCAAAAGCTGCTACTTCACTGAAAAAT	3017
QY	3129	ATGGAAGAACTGCAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTTAAAGAG	3188
Db	3018	GTGGAAGAACTGCAGGGCTCCCTGCAACCTGCAAGATGGAAGACCCCACTCTGAAAAAGAA	3077
QY	3189	AATTATCGCTTCCATGCAATCAATGGCTACATAATGGATACATCACTACCTGGCTTAGTAATG	3248
Db	3078	AATTATCGCTTCCATGCAATCAATGGCTATGATGATGATACATCACTACCTGGCTTAGTAATG	3137
QY	3249	GCTCAGGATCAAAAGGATTCGATGTAATCTGCTCAGCATGGGCAGCAATGAAAAATCCAT	3308
Db	3138	GCTCAGGATCAAAAGGATTCGATGTAATCTGCTCAGCATGGGCAGCAATGAAAAATCCAT	3197
QY	3309	TCTATTCATTTCACTGAGCAATGTCTTCACTGTACGAAAAAAAGAGAGTATATAATGGCA	3368
Db	3198	TCTATTCATTTCACTGAGCAATGTCTTCACTGTACGAAAAAAAGAGAGTATATAATGGCC	3257
QY	3369	CTGTACAATCTCTATCCAGGTGTTTGTAGACAGTGGAAATGTTAACATCCAAAGCTGGA	3428
Db	3258	CTGTACAATCTCTATCCAGGTGTTTGTAGACAGTGGAAATGTTAACATCCAAAGCTTGA	3317
QY	3429	ATTGGCGGGTGGAAATGCTTATTTGGCGAGCATCTACATGCTGGGATGAGACACACTTTT	3488
Db	3318	ATTGGCGGAATAGAAATGCTTATTTGGCGAGCATCTACATGCTGGGATGAGACACACTTTT	3377
QY	3489	CTGTGTACAGCAATAGTGTACAGACTCCCTGGGAAATGGCTTCTGGACACATTAGAGAT	3548
Db	3378	CTGTGTACAGCAATAGTGTACAGACTCCCTGGGAAATGGCTTCTGGACACATTAGAGAT	3437
QY	3549	TTTCAGATTACAGCTTCAGGACAATATGACAGTGGGGCCCCCAAAGCTGGCCAGACTTCAT	3608
Db	3438	TTTCAGATTACAGCTTCAGGACAATATGACAGTGGGGCCCCCAAAGCTGGCCAGACTTCAT	3497
QY	3609	TATTCGGGATCAATCAATGCTTGAGACACCAAGGAGCCCTTTTCTTGGATCAAGGTGAT	3668
Db	3498	TATTCGGGATCAATCAATGCTTGAGACACCAAGGATCCCCACTCTCTGATCAAGGTGAT	3557
QY	3669	CTGTGGCAACCAATGATTTATTCAGGGCATCAAGACCCAGGGGTGCCCGTCAGAACTTCC	3728
Db	3558	CTGTGGCAACCAATGATTTATTCAGGGCATCAATGACCCAGGGGTGCCCGTCAGAACTTCC	3617
QY	3729	AGCCTTACATCTCTCAGTTTATCATCATGTATAGTCTTGTATGGGAAAGAGTGGCAGACT	3788

3618 AGCCTCTACATCTCCAGTTTATCATCATGTACAGTCTTGACGGAGGAGCACTGGCAGAGT 3677
 QY 3789 TATCAGAGAAATTCACCTGGAACCTTAATGTCCTTTGGCAATGTGATTCATCTGGG 3848
 Db 3678 TACGAGGGAATTCACGGGACCTTAATGTCCTTTGGCAATGTGAGCAGCATCTGGG 3737
 QY 3849 ATAAACACAAATATTTTAAACCCCTCAATTAATGTCGATACATCCGTTTGACCCCACT 3908
 Db 3738 ATTAACACAAATATTTTAAACCCCTCAATTAATGTCGATACATCCGTTTGACCCCACT 3797
 QY 3909 CATATATAGCATTCGAGCAGCCTCTTCGATGAGTGGCTGTGATTTAAATAGTTCG 3968
 Db 3798 CATTAACAGCATCCGAGCAGCCTCTTCGATGAGTGGCTGTGATTTAAACAGTTCG 3857
 QY 3969 AGCATGCCATTGGAGATGAGAGTAAGCAATATCAGATGACAGATTAAGTCTGATCC 4028
 Db 3858 AGCATGCCCTGGAGATGAGAGTAAGCAATATCAGATGACAGATTAAGTCTGATCC 3917
 QY 4029 TACTTACCAATATGTTTCCACCTGCTCTCTCAAAAGCTGACTTCACCTCCAGGG 4088
 Db 3918 CACCTAAGCAATATATTTGCCACCTGCTCTCTCAAAAGCTGACTTCACCTCCAGGG 3977
 QY 4089 AGGAGTAATGCTGGAGACCTCAGGTGAATTAATCAAAAGAGTGGCTGCAAGTGAAGTTC 4148
 Db 3978 CGGACGAATGCTGGGAGACCTCCGGGTGAGCAGGAGAGAGTGGCTGCAAGTGAAGTTC 4037
 QY 4149 CAGAGACATGAAGTACAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4208
 Db 4038 CAGAGACATGAAGTACAGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4097
 QY 4209 ATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATGAGTGGCTCTTT 4268
 Db 4038 ATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATGAGTGGCTCTTT 4157
 QY 4269 TTTCAGATGGAAGTAAGTTCCTTCAGGGAATCAAGACTCCTTCACACCTGTGCTG 4328
 Db 4158 CTTCAGAGAGGAGCAGCAAGGTTTTCAGGGAATCAAGACTCCTTCACACCTGTGCTG 4217
 QY 4329 AACTCTCTAGACCAACCGTTACTGACCTGCTACCTTCGAATTCACCCCAAGAGTGGG 4388
 Db 4218 AACGCTCTGAGACCCCGCTGTTCACGCGCTACCTGAGGATCCACCCCAAGAGTGGG 4277
 QY 4389 CACGAGATGGCCCTGAGGATGAGGTTCTGGGCTGCGAGGACAGACCTCTACTGA 4445
 Db 4278 CAGCAGATCGCCCTGAGGCTCGAGGTTCTAGGATGTGAGGACAGAGATCTCTACTGA 4334

RESULT 14
 V23288
 V23288 standard; DNA; 4451 BP.
 DT 17-AUG-1998 (first entry)
 DE Synthetic human Factor-VIII gene lacking central B domain.
 KW Factor-VIII: blood clotting; human; synthetic gene; codon usage; ss.
 OS Homo sapiens.
 OS Synthetic.
 PN WO9812207-A1.
 PD 26-MAR-1998.
 PF 18-SEP-1997; U16639.
 PR 20-SEP-1996; US-717294.
 PA (GEHO) GEN HOSPITAL CORP.
 PI Haas J, Seed B;
 DR WPI; 98-217200/19.
 PT New synthetic eukaryotic gene(s) - in which non-preferred or less
 PT preferred codon(s) are replaced to provide high level expression in
 PT mammalian cell(s)
 PS Claim 20; Fig 13; 92pp; English.
 CC This synthetic gene codes for a human Factor-VIII protein that
 CC lacks the central B domain (amino acids 760-1639) of the native
 CC protein. In the synthetic gene, non-preferred or less preferred
 CC codons of the native gene (see V23339) are replaced by codons
 CC favored by highly expressed human genes to provide high-level
 CC expression in mammalian cells. The synthetic gene was assembled

CC from 29 pairs of oligonucleotides (see V23340-97) which served as
 CC PCR templates. Synthetic genes of the invention (see also
 CC V23289-91) are used for production of recombinant proteins in
 CC mammalian cells at levels of at least 500% of those obtained using
 CC the natural genes. They can also be used in gene therapy. An
 CC expression vector comprising a synthetic gene and a mammalian cell
 CC harboring a synthetic gene are also claimed.
 SQ Sequence 4451 BP; 972 A; 1568 C; 1234 G; 677 T;

Query Match 52.4%; Score 2531.2; DB 1; Length 4451;
 Best Local Similarity 74.1%; Pred. No. 0;
 Matches 3236; Conservative 0; Mismatches 1088; Indels 42; Gaps 1;

QY 129 GCCACCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 188
 Db 84 GCCACCCCGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 143
 QY 189 CTCGGTAGCTGCTGTGAGCAGCAAGATTTCTTCTAGAGTGCACAAATCTTTCCATTC 248
 Db 144 CTGGGCGAGCTCCCGGTGAGCAGCCCGCTTCCCGCCCGGCTGACCAAGAGCTTCCCTTC 203
 QY 249 AACACCTCAGTCGCTGACAAAGAAAGTCTGTTGTAAGATTCACGAGTACCTTTTCAAC 308
 Db 204 AACACCAAGCTGCTGACAAAGAAAGTCTGTTGTAAGATTCACGAGTACCTTTTCAAC 263
 QY 309 ATCCCTAAGCCAGAGCCACCCCTGATGGGTCTGCTAGAGTCTTCCATCCAGGCTGAGGT 368
 Db 264 ATGCGCAAGCCCGCCCGCCCTGATGGGTCTGCTAGAGTCTTCCATCCAGGCTGAGGT 323
 QY 369 TATGATCAGTGGTTCATTAATTAAGAAATGCTTCCATCTGTCAGTCTTATGCT 428
 Db 324 TACGACACCGTGGTGTATCACTTAAGAAATGCTTCCATCTGTCAGTCTTATGCT 383
 QY 429 GTTGGTATCTCTACTGAAAGCTTCTGAGGAGCTGAATATGATGATCAGACCACTCAA 488
 Db 384 GTGGGCTGAGCTACTGGAAGGCCAGGAGGCGCCGAGTACGACGACCAAGACGCTCCAG 443
 QY 489 AGGAGAAAGAAATGATTAAGTCTTCCCTGCTGAGGAGCCATATATGCTGCGAGGTC 548
 Db 444 CGCGAAGAGGAGGAGCAAGGTTTCCCGGGGGGAGCCACACCTTACGTGTGCGAGGTG 503
 QY 549 CTGAAGAAGATGCTCAATGCGCTCTGACCCACCTGCTTACCTACTCATATCTTTCT 608
 Db 504 CTTAAGAGAAAGCGCCCTATGCGCAGCGACCCCTGCTGCTGACCTACAGCTACCTGAGC 563
 QY 609 CATGTGAGCCTGTGAAGAAAGCTTAATTAAGGCTTATGAGGCGCTACTAGTATGAGA 668
 Db 564 CACGTGAGCCTGTGAAGAAAGCTTAAGGCTTATGAGGCGCTACTAGTATGAGA 623
 QY 669 GAAGGAGTCTGCGCAAGGAAAGACAGACAGACCTGCACAAAATTAATTAATTAATTTTGT 728
 Db 624 GAGGCAAGCCTGCGCAAGGAAAGACAGACAGACCTGCACAAAATTAATTAATTTTGT 683
 QY 729 GTATTGATGAAGGAAAGTGGCACTCAGAAACAAAGAACTCTGATGACAGATAGG 788
 Db 684 GTGTCAGAGAGGGAAGAGCTGCGACAGGAGACTAAGAACAGCTGATGACAGACCGC 743
 QY 789 GATGCTCATCTGCTGCGGCGCTGAGCTTAATTAAGACAGACAGTCAATGCTTATGTAACAG 848
 Db 744 GAGGCGCGCAGCGCCCGCTGAGCTTAATTAAGACAGACAGTCAATGCTTATGTAACAG 803
 QY 849 TCTGTCAGAGTCTGATGATGACAGAAATCAAGTCTATTTGATGATGATGATG 908
 Db 804 AGCTGCGCGGCTGATGATGATGACAGAAATCAAGTCTATTTGATGATGATGATG 863
 QY 909 GGCACCACTCTGAGTGCACATTAATTTCTGAGAGTCAACATTTCTTGTGAGAAC 968
 Db 864 GGCACCACTCTGAGTGCACATTAATTTCTGAGAGTCAACATTTCTTGTGAGAAC 923
 QY 969 CATGCCAGGCGCTGGAATTCGCAATATCTTCTTACTGCTCAACAGCTCTTG 1028
 Db 924 CACGCCAGGCGCTGGAATTCGCAATATCTTCTTACTGCTCAACAGCTCTTG 983

Db	3144	GCCCCCTGCACACATCCAGATGGAAGATCCACCTTCAAGAGAACTACCGCTTCCAGCC	3203
Qy	3207	ATCAATGCGCTACATAATGGATACACTACCTGGCTTAGTAATGGCTCAGATCAAGATTT	3266
Db	3204	ATCAACGGCTACATCATGTGACACACCTCGCCCGCTGGTGTATGGCCCCAGAGACCGCATC	3263
Qy	3267	CGATGGTATCTGCTCAGCATGGCGCAATGAAGAACATCCATCTATTCTTACGTGGA	3326
Db	3264	CGCTGGTACCTGCTGTCTATGGCGCAACGAGAACATCCACAGCATCCACTTCAGCGGC	3323
Qy	3327	CATGTGTTCACTGTACGAAAAAAGAGAGAGTATTAATGGCAGTGTACAATCTTATCCA	3386
Db	3324	CACGTTTTCACCGTCCGCAAGAGAGAGAGTACAGATGGCCCTGTACAACCTGTACCCC	3383
Qy	3387	GGTGTTTTGGAGACAGTGAAGTGTACCATCCAAAGCTGGAATTTGGCGGGTGAATGC	3446
Db	3384	GGCGGTTCGAGACTGTGAGATGCTGCCAGCAAGGCCGGATCTGGCGGTGGAGTGC	3443
Qy	3447	CTTATTGGCGAGCATCTACATGCTGGGATGAGCACACTTTTCTGTGTGTACAGCAATAAG	3506
Db	3444	CTGATCGGCGAGCACCTGCACGCGCGCATGAGCACCTGTCTCTGTGTGTACAGCAACAG	3503
Qy	3507	TGTACAGACTCCCTGGGAATGGCTTCTGGACACATTAGAGATTTTCAGATTACAGCTTCA	3566
Db	3504	TGCGACAGACCCCTGGGCATGGCCAGCGGCCACATCCGGCACTCCAGATACCGCCAGC	3563
Qy	3567	GGACATATGACAGTGGGCCCCCAAGCTGGCCAGACTTCAATTATCCGGATCAATCAAT	3626
Db	3564	GGCGAGTACGGCGAGTGGCTCCCAAGCTGGCCCGCTGCACATACAGCGGACGATCAAC	3623
Qy	3627	GCTTGGAGCACCAAGAGCCCTTTCTTGATCAAGGTGATCTGTGGCAACCATGATTT	3686
Db	3624	GCTTGGTGCACCAAGAGCCCTTCTCTGATCAAGGTGACCTGCTGGCCCCCATGATTC	3683
Qy	3687	ATTACGGCATCAAGAACCCAGGGTGGCCGTCAAGAGTTCTCCAGCCTCTACATCTCTCAG	3746
Db	3684	ATTCACGGCATCAAGAACCCAGGGGCCCCGCCAAGTTCAAGAGCCTGTACATCAGCCAG	3743
Qy	3747	TTTATCATCATGTATAGTCTTGATGGGAAGAGTGGCAGACTTATCGAGAAATTCCACT	3806
Db	3744	TTTATCATCATGTATAGTCTTAGACGGCAAGAGTGGCAGACTTACCGGCGCAACAGCAC	3803
Qy	3807	GGAACCTTAATGCTTCTTTGGCAATGTGATTCATCTGGGATTAACACCAATATTTT	3866
Db	3804	GGCACCTTGATGCTTCTTGGCAAGTGGACAGCAGCGGCATCAAGCAACATCTTC	3863
Qy	3867	AACCTTCAATTAATGCTGCTGATACATCCGTTGGCACCCCACTCATTAATGCAATTCCAGC	3926
Db	3864	AACCTTCAATTAATGCTGCTGATACATCCGTTGGCACCCCACTCATTAATGCAATTCCAGC	3923
Qy	3927	ACTTTCGCATGAGTGTATGGCTGTGATTAATAGTTGCAAGCATGCCATTGGGAATG	3986
Db	3924	ACCTTTCGCATGAGTGTATGGCTGTGATTAATAGTTGCAAGCATGCCATTGGGAATG	3983
Qy	3987	GAGAGTAAGCATATCAGATGCACAGATTACTGCTTCAATCTTACTTACCAATATGTTT	4046
Db	3984	GAGAGTAAGCATATCAGATGCACAGATTACTGCTTCAATCTTACTTACCAATATGTTT	4043
Qy	4047	GCCACCTGTCTCCTTCAAAAGCTGCATTCACCTCCAAAGGAGAGATATGCCCTGGAGA	4106
Db	4044	GCCACCTGTCTCCTTCAAAAGCTGCATTCACCTCCAAAGGAGAGATATGCCCTGGAGA	4103
Qy	4107	CCTCAGGTGAATATCCAAAAGAGTGGCTGCAAGTGGACTTCCAGAAGCAATGAAAGTC	4166
Db	4104	CCTCAGGTGAATATCCAAAAGAGTGGCTGCAAGTGGACTTCCAGAAGCAATGAAAGTC	4163
Qy	4167	ACAGAGTAATCTCAGGAGTAATAATCTGCTTACAGCATGTATGTGAAGAGTTT	4226
Db	4164	ACAGAGTAATCTCAGGAGTAATAATCTGCTTACAGCATGTATGTGAAGAGTTT	4223
Qy	4227	CTCATCTCCAGAGTCAAGATGGCCATCAGTGGACTCTTTTTCAGAAATGGCAAGTA	4286

Db 4224 CTGATCAGCAGCAGCCAGGACGGCCACCAGTGAACCTGTCTTCCAAAACGGCAAGTG 42833

QY 4287 AAGGTTTTACAGGGAATCAAGACTCTTACACCTGTGTGAACCTCTAGACCCACCG 4346

Db 4284 AAGGTGTTCAGAGGCAACAGACAGCTTCACACCGGTGTGAACAGCCTGACCCCCC 43433

QY 4347 TTACTGACTCGCTACCTTCGAATTCACCCCCAGAGTGGGTGCACAGATTGCCCTGAGG 4406

Db 4344 CTGCTACCCCGCTTACTGCGCATCCACCCCCAGAGCTGGGTGCACAGATCGCCCTGCGC 44033

QY 4407 ATGAGAGTCTGGGCTGCGAGGCGACAGGACCTTACTGAGGGTGGC 4452

Db 4404 ATGAGAGTCTGGGCTGCGAGGCGACAGGACCTTACTGAAAGCGGCC 4449

```

RESULT 15
ID Q50185 standard; cDNA; 9009 BP.
AC Q50185;
DT 05-MAY-1994 (first entry)
DE Human Factor VIII CDNA.
KW Blood clotting; coagulation; hybrid human/porcine factor VI
KW anti-haemophilic factor; factor 8; ss.
OS Homo sapiens.
FH Key
FT misc_feature
FT 1..2277
FT /*tag= a
FT /note= "Domain structure: equivalent to the
FT A1-A2 domain"
FT 5001..7053
FT /*tag= b
FT /note= "Domain structure: equivalent to the
FT A3-C1-C2 domain"
FT 208..7206
FT /*tag= c
FT /product= Factor_VIII
FT /note= "encodes R43257"
FT

```

PN WO9320093-A.
PD 14-OCT-1993.
PF 07-APR-1993; U03275.
PR 07-APR-1992; US-864004.
PA (UYEM-) UNITV EMORY.
PI Lollar JS, Runge MS;
DR WPI; 93-336824/42.
PT P-PSDB; R43257.
PT Hybrid human and porcine factor VIII - having high coagulant
PT activity, used for treating patients with factor VIII deficiency
PT partic. haemophilia
PS disclosure; Page 10-18; 94pp; English.
PS Hybrid human/porcine factor VIII molecules are claimed. These can be
CC obtained by (1) substituting a porcine subunit (i.e. heavy or light
CC chain) for the corresp. human subunit; (2) substituting a porcine
CC domain (i.e. A1, A2, A3, B, C1 and C2) for the corresp. human domain;
CC (3) substituting part of a porcine domain for the corresp. human
CC domain or (4) changing residues in the human sequence to the corresp.
CC porcine residues. The hybrid factor VIII molecules have greater
CC clotting activity than human factor VIII and are more resistant to
CC factor VIII inhibitors.
SQ Sequence 9009 BP; 2853 A; 1907 C; 1844 G; 2405 T;

Query Match	51.1%;	Score 2468.8;	DB 1;	Length 9009;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 2473; Conservative	0;	Mismatches	7;	Indels 0; Gaps 0;
QY 2345 AAGAAGCTTCTCCAGAACCCACCAGTCTTGAACCCCATCAACGGGAATTAACCTGTAC				
Db 5106 AAGCGTGTGCTCTCAAAACCCACCAGTCTTGAACGCCATCAACGGGAATTAACCTGTAC				
QY 2405 TACTCTTCAGTCAGATCAAGAGGAATTGACTATGATGATACCATATCAGTTGAAATGAA				
Db 5166 TACTCTTCAGTCAGATCAAGAGGAATTGACTATGATGATACCATATCAGTTGAAATGAA				

QY 2465 GAAGGAGATTTTGACATTATGATGAGATGAAAATCAGAGCCCCCGCAGCTTTCAAAA 2524
|||||
Db 5226 GAAGGAGATTTTGACATTATGATGAGATGAAAATCAGAGCCCCCGCAGCTTTCAAAA 5285
QY 2525 GAAAACAGACACTATTTTATTTGCTGCGAGTGGAGAGGCTCTGGATTATGGGATGAGTAG 2584
|||||
Db 5286 GAAAACAGACACTATTTTATTTGCTGCGAGTGGAGAGGCTCTGGATTATGGGATGAGTAG 5345
QY 2585 CTCCCCACATGTTCTAAGAAACAGGGCTGAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGT 2644
|||||
Db 5346 CTCCCCACATGTTCTAAGAAACAGGGCTGAGAGTGGCAGTGTCCCTCAGTTCAAGAAAGT 5405
QY 2645 TGTTTCCAGGAATTTACTGATGGCTCCCTTACTCAGCCCTTATACCGTGGAGAACTAAA 2704
|||||
Db 5406 TGTTTCCAGGAATTTACTGATGGCTCCCTTACTCAGCCCTTATACCGTGGAGAACTAAA 5465
QY 2705 TGAACATTTGGGACTCCCTGGGGCCATATATAGAGCAGAAAGTTGAAGATAATATCATGGT 2764
|||||
Db 5466 TGAACATTTGGGACTCCCTGGGGCCATATATAGAGCAGAAAGTTGAAGATAATATCATGGT 5525
QY 2765 AACTTTCAGAAATCAGGCTCTCGTCCCTTATTCCTTCTATTCAGCTTATTTCTATGA 2824
|||||
Db 5526 AACTTTCAGAAATCAGGCTCTCGTCCCTTATTCCTTCTATTCAGCTTATTTCTATGA 5585
QY 2825 GGAAGATCAGAGGCAAGGACAGAACTGAGAAAAAACTTGTCAAGCCTAATGAAACCA 2884
|||||
Db 5586 GGAAGATCAGAGGCAAGGACAGAACTGAGAAAAAACTTGTCAAGCCTAATGAAACCA 5645
QY 2885 AACTTACTTTTGAAAGTGCACATCATATGACCCACTAAAGATGAGTTGACTGCA 2944
|||||
Db 5646 AACTTACTTTTGAAAGTGCACATCATATGACCCACTAAAGATGAGTTGACTGCA 5705
QY 2945 AGCCTGGCTTATTTCTGATGTTGACCTGGAAGATGTGCACTCAGGCTGATTGG 3004
|||||
Db 5706 AGCCTGGCTTATTTCTGATGTTGACCTGGAAGATGTGCACTCAGGCTGATTGG 5765
QY 3005 ACCCTCTCTGCTGCGACACTAACACACAGCCCTGCTCATGGAGACAAGTGACAGT 3064
|||||
Db 5766 ACCCTCTCTGCTGCGACACTAACACACAGCCCTGCTCATGGAGACAAGTGACAGT 5825
QY 3065 ACAGGAATTTGCTCTGTTTTTCCACCATTTTGATGAGACCAAAAGCTGTACTTCACTGA 3124
|||||
Db 5826 ACAGGAATTTGCTCTGTTTTTCCACCATTTTGATGAGACCAAAAGCTGTACTTCACTGA 5885
QY 3125 AAATATGAAAAAAGAACTGAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAA 3184
|||||
Db 5886 AAATATGAAAAAAGAACTGAGGGCTCCCTGCAATATCCAGATGGAAGATCCCACTTTAA 5945
QY 3185 AGAGAAATATGCTTCATGCAATCAATGGCTTACATATGATGATACACTACTGGCTTAGT 3244
|||||
Db 5946 AGAGAAATATGCTTCATGCAATCAATGGCTTACATATGATGATACACTACTGGCTTAGT 6005
QY 3245 AATGGCTCAGATCAAAAGATTCGATGATCTGCTCAGCATGGGACGATGAAAAACAT 3304
|||||
Db 6006 AATGGCTCAGATCAAAAGATTCGATGATCTGCTCAGCATGGGACGATGAAAAACAT 6065
QY 3305 CCATTTATTTTATTTTCTAGTGAGCATGTGTTCACTGTACGAAAAAAGAGAGATATAAAT 3364
|||||
Db 6066 CCATTTATTTTATTTTCTAGTGAGCATGTGTTCACTGTACGAAAAAAGAGAGATATAAAT 6125
QY 3365 GGCACGTACATCTCTATCCAGTGTGTTTTGAGACAGTGAATGTATACCATCCAAAGC 3424
|||||
Db 6126 GGCACGTACATCTCTATCCAGTGTGTTTTGAGACAGTGAATGTATACCATCCAAAGC 6185
QY 3425 TGAATTTGGCGGTGGAATGCCTTATTTGGCGAGCATCTACATGCTGGGATGAGCACACT 3484
|||||
Db 6186 TGAATTTGGCGGTGGAATGCCTTATTTGGCGAGCATCTACATGCTGGGATGAGCACACT 6245
QY 3485 TTTCTGCTGTACAGCATAAGTGTACAGACTCCCTGGGAATGGCTTCTGACACATTTAG 3544
|||||
Db 6246 TTTCTGCTGTACAGCATAAGTGTACAGACTCCCTGGGAATGGCTTCTGACACATTTAG 6305
QY 3545 AGATTTTCAGATTACAGCTTCAAGACATATGAGACAGTGGGCCCAAGCTGGCCAGACT 3604

Db 6306 AGATTTTCAGATTACAGCTTCAGGACAAATATGAGAGTGGGCCCAAGCTGGCCAGACT 6365
QY 3605 TCATTTATCCGGATCATATCATATGCTGTGAGACACCAAGAGGCCCTTTCTTGGATCAAGT 3664
|||||
Db 6366 TCATTTATCCGGATCATATCATATGCTGTGAGACACCAAGAGGCCCTTTCTTGGATCAAGT 6425
QY 3665 GGATCTGTGGCACCAATGATTTATTCACGGGCATCAAGACCCAGGGTCCCGTCAGAAGTT 3724
|||||
Db 6426 GGATCTGTGGCACCAATGATTTATTCACGGGCATCAAGACCCAGGGTCCCGTCAGAAGTT 6485
QY 3725 CTCCAGCCTTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGAAGTGCA 3784
|||||
Db 6486 CTCCAGCCTTACATCTCTCAGTTTATCATCATGTATAGTCTTGATGGGAAGAAGTGCA 6545
QY 3785 GACTTATCGAGGAATTTCCACTGGAACCTTAATGGTCTTCTTGGCAATGTGATTCATC 3844
|||||
Db 6546 GACTTATCGAGGAATTTCCACTGGAACCTTAATGGTCTTCTTGGCAATGTGATTCATC 6605
QY 3845 TGGGATAAAAACACATATTTTAAACCCTCCAATTTATGCTCGATACATCCGTTGCACCC 3904
|||||
Db 6606 TGGGATAAAAACACATATTTTAAACCCTCCAATTTATGCTCGATACATCCGTTGCACCC 6665
QY 3905 AACTCATTTATAGCATTCGACAGCACTCTTCGCATGAGTTGATGGGCTGATTTAAATAG 3964
|||||
Db 6666 AACTCATTTATAGCATTCGACAGCACTCTTCGCATGAGTTGATGGGCTGATTTAAATAG 6725
QY 3965 TTGCAGCATGCCATTTGGGAATGGAGAGTAAAGCAATATCAGATGCAAGATTACTGCTTC 4024
|||||
Db 6726 TTGCAGCATGCCATTTGGGAATGGAGAGTAAAGCAATATCAGATGCAAGATTACTGCTTC 6785
QY 4025 ATCTTACTTTTACCAATATGTTTGCCACCTGCTCTCTCAAAAAGCTGACTTACCTCCA 4084
|||||
Db 6786 ATCTTACTTTTACCAATATGTTTGCCACCTGCTCTCTCTCAAAAAGCTGACTTACCTCCA 6845
QY 4085 AGGAGAGTATATGCTCTGAGAGACCTTCAGGTGAATATCCAAAAGAGTGGCTGCAAGTGA 4144
|||||
Db 6846 AGGAGAGTATATGCTCTGAGAGACCTTCAGGTGAATATCCAAAAGAGTGGCTGCAAGTGA 6905
QY 4145 CTTCAGAGACAAATGAAAGTCAACAGGAGTAACTACTCAGGAGTAAATCTCTGCTTAC 4204
|||||
Db 6906 CTTCAGAGACAAATGAAAGTCAACAGGAGTAACTACTCAGGAGTAAATCTCTGCTTAC 6965
QY 4205 CAGCATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCT 4264
|||||
Db 6966 CAGCATGTATGTGAAGAGTTCCTCATCTCCAGCAGTCAAGATGGCCATCAGTGGACTCT 7025
QY 4265 CTTTTTTCAGAAATGGCAAGTAAAGTTTTTTCAGGGAATCAAGACTCCTTACACCTGT 4324
|||||
Db 7026 CTTTTTTCAGAAATGGCAAGTAAAGTTTTTTCAGGGAATCAAGACTCCTTACACCTGT 7085
QY 4325 GGTGAACCTCTAGACCCACCGTTACTGACTCGTACCTTCGAATTCACCCCAAGAGTTG 4384
|||||
Db 7086 GGTGAACCTCTAGACCCACCGTTACTGACTCGTACCTTCGAATTCACCCCAAGAGTTG 7145
QY 4385 GGTGACACGATTTGCCCTGAGAGTGAAGTTCTGGCTGCGAGGACAGGACCTTACTG 4444
|||||
Db 7146 GGTGACACGATTTGCCCTGAGAGTGAAGTTCTGGCTGCGAGGACAGGACCTTACTG 7205
QY 4445 AGGTTGGCCACTGACACACTGCACTGCCGTGCACTCTCCCTCCAGCTCCAGGGCAG 4504
|||||
Db 7206 AGGTTGGCCACTGACACACTGCACTGCCGTGCACTCTCCCTCCAGCTCCAGGGCAG 7265
QY 4505 TGTCCCTCCCTGGCTTGCTTCTTACCTTTGTGCTAAATCCTAGACACACTGCCTGAAG 4564
|||||
Db 7266 TGTCCCTCCCTGGCTTGCTTCTTACCTTTGTGCTAAATCCTAGACACACTGCCTGAAG 7325
QY 4565 CCTCTGAATTAACATATCATCAGTCTGCATTTCTTTGGTGGGGGCCAGAGGCTGCAT 4624
|||||
Db 7326 CCTCTGAATTAACATATCATCAGTCTGCATTTCTTTGGTGGGGGCCAGAGGCTGCAT 7385
QY 4625 CCAATTTAACTTAACCTTACCTAATTTTCTGACGCTGCTCCAGATTAATCTCCTTCC 4684
|||||

Db 7386 CCAATTTAACTTAACCTCTTACCCTATTTCTGACAGCTGCTCCAGATTAATCTCTTCCTTCC 7445
QY 4685 AATATACTAGGCAAAAAGAGTGAGAGAAACCTGCATGAAAGCATCTTCCCTGAAAA 4744
|||
Db 7446 AATATACTAGGCAAAAAGAGTGAGAGAAACCTGCATGAAAGCATCTTCCCTGAAAA 7505
QY 4745 GTTAGGCTCTCTCAGAGTCACCACTTCCTCTGTGTAGAAAAAACTATGTGATGAACCTTG 4804
|||
Db 7506 GTTAGGCTCTCTCAGAGTCACCACTTCCTCTGTGTAGAAAAAACTATGTGATGAACCTTG 7565
QY 4805 AAAAGATATTATGATGTT 4824
|||
Db 7566 AAAAGATATTATGATGTT 7585

Search completed: August 18, 1999, 14:52:44
Job time: 1392 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

CM nucletic - nucletic search, using sw model

Run on: August 18, 1999, 14:29:31 ; Search time 758.83 Seconds
(without alignments)
12560.497 Million cell updates/sec

Title: US-09-001-039a-46
Perfect score: 4832
Sequence: 1 CTCGAGCTAAAGATATTTA.....ATTATGATGTGCGCCGC 4832

Scoring table: IDENTITY_NUC
Searched: 2546578 seqs, 986266752 residues

Database :
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: em_est20:*
21: em_est21:*
22: em_est22:*
23: em_est23:*
24: em_est24:*
25: em_est25:*
26: em_est26:*
27: em_est27:*
28: em_est28:*
29: em_est29:*
30: em_est30:*
31: em_est31:*
32: em_est32:*
33: em_est33:*
34: em_est34:*
35: em_est35:*
36: em_est36:*
37: em_est37:*
38: em_est38:*
39: em_est39:*
40: em_est40:*
41: em_est41:*
42: em_est42:*
43: em_est43:*
44: em_est44:*
45: em_est45:*
46: em_est46:*
47: em_est47:*
48: em_est48:*
49: em_est49:*
50: em_est50:*
51: em_est51:*
52: em_est52:*
53: em_est53:*

54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	347.4	7.2	473	22	R51152	R51152 yg71e07.r1
2	339	7.0	431	21	T74128	T74128 yc82a11.r1
3	299.2	6.2	559	23	H16894	H16894 ym39a03.r1
4	290	6.0	291	21	F12385	F12385 HSC39C111.n
5	263.8	5.5	517	43	A1176907	A1176907 EST220513
6	239.8	5.0	505	46	A1410862	A1410862 EST239155
7	201.4	4.2	422	45	A1233991	A1233991 EST230679
8	200	4.1	400	48	A1602045	A1602045 UI-R-G0-u
9	194	4.0	265	29	AA184901	AA184901 mt93e12.r
10	152	3.1	497	41	A1072568	A1072568 UI-R-C2-n
11	131.8	2.7	864	45	A1324553	A1324553 mr92d06.y
12	130	2.7	657	43	AA1173222	AA1173222 ud63a12.x
13	121.6	2.5	423	40	AA984429	AA984429 am86c08.s
14	121.4	2.5	633	30	AA208846	AA208846 mw75a01.r
15	120.8	2.5	1006	43	A1225600	A1225600 u106c05.y
16	120.6	2.5	613	34	AA512671	AA512671 vg39g02.r
17	117	2.4	456	45	AA817854	AA817854 UI-R-A0-a
18	114.8	2.4	785	43	A1181867	A1181867 ud63a12.y
19	110	2.3	518	39	AA839539	AA839539 uc99a11.r
20	105.6	2.2	521	39	AA865338	AA865338 og88d10.s
21	104.8	2.2	864	48	A1573566	A1573566 ud45f11.y
22	102.6	2.1	365	26	W45868	W45868 mc81h11.r1
23	101.8	2.1	625	47	A1526753	A1526753 u142e03.y
24	100.4	2.1	859	45	A1326571	A1326571 mm20g07.y
25	98.6	2.0	629	50	A1683839	A1683839 tx77g04.x
26	97.2	2.0	427	38	AA771447	AA771447 vm45b04.r
27	96	2.0	567	39	AA873581	AA873581 oh68a12.s
28	94.8	2.0	462	29	AA166973	AA166973 z087e05.r
29	94.2	1.9	206	48	A1547643	A1547643 UI-R-C3-s
30	93.8	1.9	477	34	AA461838	AA461838 vf95d04.r
31	93.8	1.9	315	48	A1611296	A1611296 tui6h02.x
32	92.6	1.9	667	28	AA106138	AA106138 mm20g07.r
33	91.8	1.9	586	26	W83411	W83411 mf28d06.r1
34	91.4	1.9	689	43	A1195467	A1195467 u152c06.x
35	90.2	1.9	820	43	A1194959	A1194959 u158h03.x
36	90	1.9	372	21	T61867	T61867 yb93a02.s1
37	87	1.8	800	20	T28497	T28497 EST46762.Hu
38	87	1.8	800	42	A1114892	A1114892 u149f07.x
39	86.6	1.8	530	43	A1195667	A1195667 u152c06.y
40	86	1.8	547	30	AA238246	AA238246 mx76d03.r
41	86	1.8	685	49	A1646857	A1646857 ub66f09.x
42	85.8	1.8	489	30	AA270973	AA270973 va84c09.r
43	85	1.8	475	44	A1265337	A1265337 uk01h06.y
44	84.8	1.8	392	30	AA221925	AA221925 mw13e10.r
45	84.6	1.8	666	44	A1323769	A1323769 mm20g07.x

ALIGNMENTS

RESULT 1
R51152
LOCUS R51152 473 bp mRNA
DEFINITION yg71e07.r1 Soares infant brain INIB Homo sapiens CDNA clone
IMAGE:38551 5' similar to gb:M1413 COAGULATION FACTOR VIII
PRECUSOR (HUMAN)'' mRNA sequence.
ACCESSION R51152
NID 9813054


```

VERSION      R51152.1  GI:813054
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
AUTHORS      Eutheria; Primates; Catarrhini; Hominidae; Homo.
              Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
              Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
              Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
              Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
              Wilson, R.
TITLE        The WashU-Merck EST Project
JOURNAL      Unpublished (1995)
COMMENT      Contact: Wilson RK
              Washington University School of Medicine
              4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
              Tel: 314 286 1800
              Fax: 314 286 1810
              Email: est@watson.wustl.edu
              High quality sequence stops: 351
              Source: IMAGE Consortium, LNL
              This clone is available royalty-free through LNL; contact the
              IMAGE Consortium (info@image.lnl.gov) for further information.
              Seq primer: M13RP1
              High quality sequence stop: 351.
FEATURES
source
1. 473
/organism="Homo sapiens"
/db_xref="GDB:411092"
/db_xref="taxon:9606"
/clone="IMAGE:38551"
/clone_lib="Soares infant brain INIB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lfamid BA; Site_1: Not
I; Site_2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5';
AACTGGAAGAATTCGCGCGCAGGAATTTTTTTTTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lfamid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
COUNT      116 a 118 c 104 g 132 t 3 others
IN
Query Match      7.2%; Score 347.4; DB 22; Length 473;
Best Local Similarity 96.8%; Pred. No. 1.3e-88;
Matches 419; Conservative 0; Mismatches 6; Indels 8; Gaps 6;
QY 4393 AGATTGCCCTGAGATGAGAGTTCGGGCTGCGAGCACAGACCTCTACTGAGGTGGC 4452
      |||||||
      1 AGATTGCCCTGAGATGAGAGTTCGGGCTGCGAGC-ACAGACCTCTACTGAGGTGGC 59
QY 4453 CACTGCAGACCTGCCACTGCCGTCACCTCTCCCTCCTCAGCTCCAGGGCAGTGTCCCTC 4512
      |||||||
      60 CACTGCAGACCTGCCACTGCCGTCACCTCTCCCTCCTCAGCTCCAGGGCAGTGTCCCTC 119
QY 4513 CCTGGCTTGCCTTACCTTTGTGCTAAATCCTAGACAGACACTGCTTGAAGCCTCCTGA 4572
      |||||||
      120 CCTGGCTTGCCTTACCTTTGTGCTAAATCCTAGACAGACACTGCTTGAAGCCTCCTGA 178
QY 4573 ATTAACTATCATCAGTCTGTCATTTCTTTGGTGGGGGCCAGAGAGGTGCATCCAATTGA 4632
      |||||||
      179 ATTAACTATCATCAGTCTGTCATTTCTTTGGTGGGGGCCAGAGAGGTGCATCCAATTGA 238
QY 4633 ACTTAACCTTACCTATTCTTCTGACAGCTGCTCCAGATTACTCCTTCTCCAATATAAC 4692
      |||||||
      239 ACTTAACCTTACCTATTCTTCTGACAGCTGCTCCAGATTACTCCTTCTCCAATATAAC 298

```

```

QY 4693 TAGGCAAAAAGAGTGAGAGAAACCT-GCATGAAGCATTTCTCCCTGAAAGTTAGGC 4751
      |||||||
      299 TAGGCAAAAAGAGTGAGAGAAACCTGGCATGAAGCATTTCTCCCTGAAAGTTAGGC 358
QY 4752 CTCTCAGAGTCACCACT-TCCTCTGTGTGA-GAAAACCTATGTG---ATGAACCTTGAA 4806
      |||||||
      359 CTCTCAGAGTCACCACTTTCCTCTGTGTGAAGAAAACCTATGTGATGAACCTTGAA 418
QY 4807 AAAGATATTATG 4819
      |||||||
      419 AAAGATATTATG 431

```

```

RESULT      2
T74128      431 bp  mRNA  EST 07-MAR-1995
LOCUS       yc82a11.r1 Soares infant brain INIB Homo sapiens cDNA clone
DEFINITION  IMAGE:22304 5' similar to gb:M14113 COAGULATION FACTOR VIII
            PRECURSOR (HUMAN); mRNA sequence.
ACCESSION   T74128
NID         9690803
VERSION     T74128.1 GI:690803
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 431)
AUTHORS     Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
            Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
            Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
            Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
            Wilson, R.
TITLE       The WashU-Merck EST Project
JOURNAL     Unpublished (1995)
COMMENT

```

```

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 4827
High quality sequence stops: 373 Source: IMAGE Consortium, LNL This
clone is available royalty-free through LNL; contact the IMAGE
Consortium (info@image.lnl.gov) for further information.
Insert Length: 4827 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 373.
FEATURES
source
1. 431
/organism="Homo sapiens"
/db_xref="GDB:394651"
/db_xref="taxon:9606"
/clone="IMAGE:22304"
/clone_lib="Soares infant brain INIB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lfamid BA; Site_1: Not
I; Site_2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5';
AACTGGAAGAATTCGCGCGCAGGAATTTTTTTTTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lfamid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT  110 a 110 c 84 g 120 t 7 others
ORIGIN

```

Query Match 7.0%; Score 339; DB 21; Length 431;
Best Local Similarity 97.7%; Pred. No. 3.2e-86;
Matches 383; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

QY 4430 ACAGACCTCTACTGAGGGTGGCCACTGACGACCACTGCCACTGCC-CTCACCCTCTCCCTC 4488
|||||
DB 1 ACAGGACCTCTACTGAGGGTGGCCACTGACGACCACTGCCACTGCCCTC 60
QY 4489 CTCAGCTCCAGGGCAGTGTCCCTCCCTGGCTTGCCTTCTACCTTGTGTAAATCCTAGC 4548
|||||
DB 61 CTCAGCTCCAGGGCAGTGT-CTTCCCTGGCTTG-CTTCTACCTTGTGTAAATCCTAGC 118
QY 4549 AGACACTGCTTGAAGCCTCTGAATTAATCATCATGCTCTGCAATTTCTTGTGGGG 4608
|||||
DB 119 AGACACTGCTTGAAGCCTCTGAATTAATCATCATGCTCTGCAATTTCTTGTGGGG 178
QY 4609 GGGCAGGAGGGTGCATCCAACTTAATTAATCTTACCTATTTTTCGAGCTGCTCCAG 4668
|||||
DB 179 GGGCAGGAGGGTGCATCCAACTTAATTAATCTTACCTATTTTTCGAGCTGCTCCAG 238
QY 4669 ATTACTCTCTCTCCATATATTAATGAGCAAAAAGAGTGAAGAACTGCATGAAG 4728
|||||
DB 239 ATTACTCTCTCTCCATATATTAATGAGCAAAAAGAGTGAAGAACTGCATGAAG 298
QY 4729 CATCTTCCCTGAAAGTTAGCGCTCTGAGAGTCAACCACTCTCTGTGTGA-GAAAAAC 4787
|||||
DB 299 CATCTTCCCTGAAAGTTAGGNCCTGAGAGTCAACCACTCTCTGTGTGAGANAAN 358
QY 4788 TATGTATGAACCTTGAAGAAATATTATG 4819
|||||
DB 359 TATGTATGAACCTTGAAGAAATATTATG 390

RESULT 3
H16894 559 bp mRNA EST 29-JUN-1995
LOCUS ym39a03.r1 Soares infant brain INIB Homo sapiens cDNA clone
DEFINITION IMAGE:50454 5' similar to gb:M14113 COAGULATION FACTOR VIII
PRECURSOR (HUMAN);, mRNA sequence.
H16894
9883134
ACCESSION H16894.1 GI:883134
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 559)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT On May 18, 1995 this sequence version replaced.

TITLE
JOURNAL
COMMENT

FEATURES
Source 1. .559
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="GDB:423264"
/db_xref="taxon:9606"
/clone="IMAGE:50454"
/clone_lib="Soares infant brain INIB"
/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: Lactid BA; Site_1: Not
I; Site_2: Hind III; 1st strand cDNA was primed with a Not
I - o1190(dt) primer [5]
ACTGGAAGATTCGGCGCCGACAGCAATTTTCTTTTCTTTT 3';
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the Lactid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

BASE COUNT 131 a 130 c 126 g 162 t 10 others
ORIGIN

Query Match 6.2%; Score 299.2; DB 23; Length 559;
Best Local Similarity 95.2%; Pred. No. 9.4e-75;
Matches 374; Conservative 0; Mismatches 9; Indels 10; Gaps 6;

QY 4393 AGATTGCCCTGAGGATGAGGTTCTGGGCTGCGAGGACAGGACCTCTACTGAGGTGGC 4452
|||||
DB 1 AGATTGCCCTGAGGATGAGGTTCTGGGCTGCGAGG-ACAGGACCTCTACTGAGGTGGC 59
QY 4453 CACTGCAGACCTGCCACTGCCGTACCTCTCCCTCTCAGCTCCAGGCGAGTGCCTC 4512
|||||
DB 60 CACTGCAGACCTGCCACTGCCGTACCTCTCCCTCTCAGCTCCAGGCGAGTGCCTC 119
QY 4513 CCTGGCTTGCCTTCTACTTTGTGCTAAATCCTAGACAGACTGCCCTGAAGCCTCTGA 4572
|||||
DB 120 CCTGGCTTGCCTTCTACTTTGTGCTAAATCCTAGACAGACTGCCCTGAAGCCTCTGA 178
QY 4573 ATTAACATCATCAGTCTGCAATTTCTTGTGGGGGCGGAGGAGGTCATCCAATTA 4632
|||||
DB 179 ATTAACATCATCAGTCTGCAATTTCTTGTGGGGGCGGAGGAGGTCATCCAATTA 238
QY 4633 ACTTAACCTTACCTATTTTCTGACAGCTGCCAGATTAATCTCTTCCATATTAAC 4692
|||||
DB 239 ACTTAACCTTACCTATTTTCTGACAGCTGCCAGATTAATCTCTTCCATATTAAC 298
QY 4693 TA-GGCAAAAAGAGTGAAGGAGAAACCTGCAT-GAAACATTTCTT--CCCTGAAAAGTTA 4748
|||||
DB 299 TAGGCAAAAAGAGTGAAGGAGAAACCTGCATGAAACATTTCTTCCCTGAAAAGTTA 358
QY 4749 GGGCTCTCAGAG---TCACCACCTTCTCTGT 4777
|||||
DB 359 GGGCTTTCAGAGTTCACCACTTCTCTGT 391

RESULT 4
F12385 291 bp mRNA EST 13-MAR-1995
LOCUS HSC39C111 normalized infant brain cDNA Homo sapiens cDNA clone
DEFINITION c-39c11, mRNA sequence.
F12385
9706734
ACCESSION F12385.1 GI:706734
VERSION EST.
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 291)
AUTHORS Auftray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C.,
Devignes,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B.,
Lorenzo,F., Mitchell,H., Mariage-Samson,R., Pletcu,G., Pouliot,Y.,
Sebastiani-Kabekchis,C. and Tessier,A.
TITLE IMAGE: molecular integration of the analysis of the human genome

and its expression
JOURNAL C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
MEDLINE 95277534
COMMENT

Contact: Genethon
Genexpress-Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read.
Genexpress_library_idt: C; Genexpress_sequence_idt: y1c-39c11
Insert Length: 1570 Std Error: 0.00
Seq primer: (-21)M13-universal
High quality sequence stop: 417.
Location/Qualifiers

FEATURES

source

1. 291
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="21"
/clone="c-39c11"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lafmid BA; Site_1: HindIII;
Site_2: NotI; sex=Female; dev_stage=3 months old;
isolate=muscular atrophy patient; tissue_type=total
brain; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B.Soaes, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"
BASE COUNT 52 a 92 c 66 g 80 t 1 others
ORIGIN

Query Match 6.0%; Score 290; DB 21; Length 291;
Best Local Similarity 99.7%; Pred. No. 3e-72;
Matches 290; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4391 CCAGATTGCCCTGAGGATGAGTTCCTGGCTGCGAGGACAGACCTTACTGAGGTG 4450
1 CCAGATTGCCCTGAGGATGAGTTCCTGGCTGCGAGGACAGACCTTACTGAGGTG 60
4451 GCCACTGAGCACCTGCCACTGCCGTGACCTCTCCCTCAGCTCCAGGGCAGTGTCC 4510
61 GCCACTGAGCACCTGCCACTGCCGTGACCTCTCCCTCAGCTCCAGGGCAGTGTCC 120
QY 4511 TCCCTGGCTGCTTCTACCTTTGTGCTAAATCCTAGACACACTGCTTGAAGCTCCT 4570
121 TCCCTGGCTGCTTCTACCTTTGTGCTAAATCCTAGACACACTGCTTGAAGCTCCT 180
DB 4571 GAATTAATCATCATCATCTGATTTCTTTGGTGGGGGCCAGAGGTCATCCAATT 4630
181 GAATTAATCATCATCATCTGATTTCTTTGGTGGGGGCCAGAGGTCATCCAATT 240
QY 4631 TAATTAATCTTACTATTTCTTGACAGCTGCTCCAGATTACTCCTTCT 4681
241 TAATTAATCTTACTATTTCTTGACAGCTGCTCCAGATTACTCCTTCT 291

RESULT 5
A1176907/c 517 bp mRNA EST 20-JAN-1999
LOCUS A1176907
DEFINITION EST220513 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone
ROVBY42 3' end, mRNA sequence.
ACCESSION A1176907
NID 93727545
VERSION A1176907.1 GI:3727545
KEYWORDS EST.
SOURCE Rattus sp.

ORGANISM

Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 517)
AUTHORS Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J.,
Kerlavage, A.R. and Adams, M.D.
TITLE Rat Genome Project: Generation of a Rat EST (REEST) Catalog & Rat
Gene Index
JOURNAL Unpublished (1998)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2152025.

Contact: Lee, NH
ATCC
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.
Location/Qualifiers

FEATURES

source

1. 517
/organism="Rattus sp."
/db_xref="ATCC (inhost):2031594"
/db_xref="taxon:10118"
/clone="ROVBY42"
/note="Organ: ovary; Vector: pT73Pac; Site_1: EcoRI;
Site_2: NotI"
BASE COUNT 93 a 174 c 154 g 96 t
ORIGIN

Query Match 5.5%; Score 263.8; DB 43; Length 517;
Best Local Similarity 70.9%; Pred. No. 1.3e-64;
Matches 365; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

QY 2502 CAGAGCCCCCGCAGCTTCAAAAGAAAACAGACACTATTATTGCTGCAGTGGAGAG 2561
DB 513 CAGGACCCCGCGGCTTTCAGCAGAAAACCGGCACTATTTCATGCGCCCGTGAAGAG 454
QY 2562 CTCTGGATTATGGGATGAGTACCTCCACATGTCTTAAGAACAGGGCTCAGAGTGGC 2621
DB 453 CTTTGGCACTACGGGCTGCGTGTGTCCCTGAGCTCCGGAGACAGGGCGTGAAGTGG 394
QY 2622 AGTGTCCCTCAGTTAAGAAAGTTTTCAGGAATTTACTGATGGCTCCTTACTCAG 2681
DB 393 GACGACAGCTAGGTTCCGGAAGTGTTCCTCGTACGTTACAGGAGGGGTGTCAGCGAG 334
QY 2682 CCCTTATACCGTGGAGAACTAATATGACATTTGGGACTCCTGGGGCATATATAAGACA 2741
DB 333 CGCGTGAACCGTGGGAGCTGGAGCACACCTGGGGCTACTGGGGCATATACATACGGGCG 274
QY 2742 GAAGTTGAAGATATATCATGTGTAATTTTCAGAAATCAGGCTCTGCTTCCCTATTCCTTC 2801
DB 273 GAAGTGAAGGACACATCGTGTGACCTTCGGGAACAGGCTCCCGCCCTACTCTTC 214
QY 2802 TATTCTAGCCTTATTTCTTATGAGGAAGATCAGAGGCAAGAGACAGAACTAGAAAAAC 2861
DB 213 TACTTCAGCCTGTGTCTTACCCGGAAGCCGAGCGT--GGAGCTGCACCAAGAGCAAC 157
QY 2862 TTTGTCAAGCCTAATGAACCAAACTTACTTTTGAAGAGTGAACATCATATAGCAACC 2921
DB 156 TTTGTGCGGGCCCAACGAGACCAAACTTACTTGTGAAGAGTCCGGCCCCACATAGCGGCC 97
QY 2922 ACTAAGATGAGTTTGAAGTGAAGGCTGGGCTTATTTCTGTGATGTGACCTGGAAAAA 2981
DB 96 ACGGACGGCGAGTTTGAAGTGAAGGCTGGGCTTATTTCTCGACGTGAGACCTGAGCGGA 37
QY 2982 GATGTGCACTCAGGCTGATTTGAGCCCTTCTGCT 3016
DB 36 GACCTGCACTCGGGGCTGATTTGGGCCACTGCTGCT 2

[illegible]

QY	1226	TTCCTTTATCCAAATTGCGCTCAGTGTGCCAAGAAGCATCCTAAACTGGGTACATTACAT	1285
Db	92	--CCCTTCGTCACAGGTCCGCTCAGACAGCCCAAGAGACACCACAAGACTTGGTGCACATACAT	35
QY	1286	TGCTGCTGAAGAGAGGAGCTGGGACTATGCTCC	1318
Db	34	CGCAGCAGAGAGAGGGGACTGGGACTACGCGCC	2
RESULT	7		
LOCUS	AI233991		
DEFINITION	AI233991	422 bp mRNA EST	08-JAN-1999
ACCESSION	ESTJ30679	Normalized rat lung, Bento Soares Rattus sp.	cDNA clone
NID	RUCR80 3'	end, mRNA sequence.	
VERSION	AI233991		
KEYWORDS	g4135942		
SOURCE	AI233991.1	GI:4135942	
ORGANISM	EST.		
REFERENCE	Rattus sp.		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;		
JOURNAL	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
COMMENT	1 (bases 1 to 422)		
	Lee,N.H., Glodok,A., Chandra,I., Mason,T.M., Quackenbush,J.,		
	Kerlavage,A.R. and Adams,M.D.		
	Rat Genome Project: Generation of a Rat Est (REST) Catalog & Rat		
	Gene Index		
	Unpublished (1998)		
	On Oct 30, 1998 this sequence version replaced gi:3817871.		
FEATURES	Contact: Lee, NH		
source	ATCC		
	The Institute for Genomic Research		
	9712, Medical Center Drive, Rockville, MD 20850, USA		
	Tel: (301)-838-3529		
	Fax: (301)-838-0208		
	Email: nhlee@igr.org		
	Seq primer: M13-21.		
	Location/Qualifiers		
	1..422		
	/organism="Rattus sp."		
	/db_xref="taxon:10118"		
	/clone="RUCR80"		
	/clone_lib="Normalized rat lung, Bento Soares"		
	/note="Organ: lung; Vector: pT7T3pac; Site_1: EcoRI;		
	Site_2: NotI"		
BASE COUNT	74 a 120 c 160 g 68 t		
ORIGIN			
Query Match	4.2%; Score 201.4; DB 45; Length 422;		
Best Local Similarity	67.5%; Pred. No. 9e-47;		
Matches 283; Conservative	0; Mismatches 136; Indels 0; Gaps 0;		
QY	3061	CAGTACAGGAATTGCTCTGTTTTCACCATCTTGATGAGACCAAAGCTGGTACTTCA	3120
Db	3	CAGTGCAGGAGTTCGCGCTGCTCTTCGCGGTTCGATGAGACCAAGAGCTGTACTGCG	62
QY	3121	CTGAATAATGGAAGAACTGCAGGGGCTCCCTGCAATATCCAGATGGAAGATCCCCTT	3180
Db	63	CTGAGAACCTTGACCGGGAAGTGACAGGCCGCCCGCGGCACCCCGCGGAGACCCGCT	122
QY	3181	TTAAGAGAATTATGCTCTTCATGCAATCAATGGCTACATAATGATACACTACTGGCT	3240
Db	123	GGAGAGAGAGTAGTACCGCTTCATATCGCGTGAATGTTACGTCGACAGATGCTCTGCTTGCC	182
QY	3241	TAGTAATGGCTCAGAGTCAAGGATTCGATGGTATCTGCTCAGAGTGGGAGCAATGAAA	3300
Db	183	TGGTGTGCGCGGAAGTCCGAGGAGCAGAGGTGGCATCTGCTCAGCATGGCGGACCCGCGC	242
QY	3301	ACATCCATCTATTCATTTCAGTGGACATGTGTTCACTGTACGAAAAAAGAGAGATATA	3360

DB	ACC	CHROM	START	END	ORIGIN	FEATURE	COORDINATE	SEQUENCE
DB	243	ACGCACAGTCCGTCGATTTTCAGCGGCGCACAGCGTCACCGCTGGCGGAGCGGGCGGAGCACA	302					
QY	3361	AAATGGCACTGTACAATCTCTATCCAGGTGTTTGGAGACAGTGGAAAATGTACCATCCA	3420					
DB	303	GGACGGCGGTGTGCAACCTGTACCCAGGCGTGTTCACAAACCGTGAGATGTGCCGTCCA	362					
QY	3421	AAGCTGCAATTTGGCGGGGTGGAATGCCCTTATTGGCGAGCATCTACATGCTGGATGAC	3479					
DB	363	GAGCGGGGATCTGGCGGGGTGGAGAGCGCTGTGGCGAGACCACTACGCGCTGGGATGAC	421					
RESULT	8							
LOCUS	AI602045/c							
DEFINITION	AI602045	400 bp	mRNA	EST	21-APR-1999			
ACCESSION	AI602045	UI-R-G0-us-d-09-0-UI.s2	UI-R-G0	Rattus norvegicus	cdna clone			
VERSION	94611206	AI602045.1	GI:4611206					
KEYWORDS	EST.							
SOURCE	Norway rat.							
ORGANISM	Rattus norvegicus							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;							
AUTHORS	Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.							
TITLE	1 (bases 1 to 400)							
JOURNAL	Bonaldi,M.F., Lennon,G. and Soares,M.B.							
MEDLINE	Normalization and subtraction: two approaches to facilitate gene							
COMMENT	discovery							
	Genome Res. 6 (9), 791-806 (1996)							
	97044477							
	On May 7, 1998 this sequence version replaced gi:3121068.							

Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Oligo-dT track not found. Not I site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: clones will be available through
 Research Genetics (www.resgen.com)
 Seq primer: M13 Forward.

FEATURES
 source
 Location/Qualifiers
 1..400
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-G0-us-d-09-0-UI"
 /clone_lib="UI-R-G0"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pTR3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-G0
 library is a normalized library constructed from a
 mixture of rat tissues (nodose ganglia, dorsal root
 ganglia, and trigeminal ganglia). The tag is a string of
 6 nucleotides present between the Not I site and the
 oligo-dT track. The library was constructed as described
 by Bonaldo, Lennon and Soares, Genome Research 6:
 791-806, 1996."

BASE COUNT
 ORIGIN
 71 a 120 c 134 g 75 t

[illegible]

QY 2715 GGACTCTGGGGGCATATATTAAGAGCAGAGTGAAGATATATCATGGTAACTTTCAGA 2774
 || || ||||| ||||| || || ||||| || || ||||| || || ||||| || || |||||
 Db 339 GGGCTACTGGGGCCATACTTACGGGGCGAAGTGAAGACAAATCGTGGTGAACCTTCCGG 280
 QY 2775 AATCAGGCTCTCGTCCCTATTCCTTAATTCCTAGCCTTAATTTCTTATGAGGAAGATCAG 2834
 || ||||| ||||| || || ||||| || || ||||| || || ||||| || |||||
 Db 279 AACCAGGCTCCCGCCCATACTCATTAATCTCCAGCCTCGTGTCTTACC CGGAAGCCGAC 220
 QY 2835 AGGCAAGGAGCAGAACCTAGAAAAAACTTTGTCAAGCCTAATGAACCAAACTTACTTT 2894
 | ||||| || || || ||||| || || ||||| || || ||||| |||||
 Db 219 GG--TGAGCTGCACCAAGAGCAACTTTGTGCGGCCCAACGAGACCAAACTTACTTC 163
 QY 2895 TGAAGAAGTCAACATCATATGCGCACCCACTAAAGATGAGTTGACTGCAAGCCTGGGCT 2954
 ||||| || || ||||| ||||| || ||||| ||||| ||||| ||||| |||||
 Db 162 TGAAGAGTCCGGCCCAACATGGCGCCACGAGACGGCGAGTTGACTGCAAGCCTGGGCT 103
 QY 2955 TATTTCTGTATGTGACCTGGA AAAAGATGTGCACCTCAGGCCCTGATTTGACCCCTTCTG 3014
 ||||| || || ||||| || || ||||| || ||||| || ||||| || || |||||
 Db 102 TATTTCTCCGACGTGACCTCGAGCGAGACCTGCACCTCGGGGCTGATTTGGGCCACTGCTC 43
 QY 3015 GTCTGCCACACTAACACACTGAACCTGCTCATGGGAGAC 3054
 || ||||| || || ||||| || || ||||| || || ||||| || || |||||
 Db 42 GTGTGCCACGCTCCACCCCTGCACCCCTCGCACGCGCGTC 3

RESULT 9
AA184901
LOCUS
DEFINITION
AA184901 265 bp mRNA EST 17-FEB-1997
m33e12.r1 Soares mouse 3NBMS Mus musculus cDNA clone IMAGE:637486
5' similar to gb:L05573 Mus domesticus coagulation factor VIII
mRNA, complete (MOUSE);, mRNA sequence.
AA184901
ACCESSION
NID g1768759
VERSION AA184901.1 GI:1768759
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 265)
AUTHORS Maria,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1292355.

Contact: Marra M/Mouse EST Project
 Washu-HHMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:389478
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 250.
 Location/Qualifiers
 1. .265
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:637486"
 /clone_lib="Soares mouse 3NbMs"
 /sex="male"
 /tissue_type="Spleen"
 /dev_stage="4 weeks"
 /lab_host="DH10B"

NID
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

g4058982
 AI324553.1 GI:4058982
 EST.
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 864)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Gelsel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 The WashU-HHMI Mouse EST Project
 Unpublished (1996)
 On Jan 14, 1998 this sequence version replaced gi:1796937.

Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:370339

This read is a RESEQUENCE of a previously sequenced mouse clone
This read has been verified (found to hit its original self in the
correct orientation)
Seq primer: -40RP from Gibco
High quality sequence stop: 418.

FEATURES	Location/Qualifiers
source	1. .864

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:604907"
/clone_lib="Stratagene mouse embryonic carcinoma
(#937317)"
/tissue_type="carcinoma"
/dev_stage="embryonic"
/lab_host="SOLR (kanamycin resistant)"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. p19 cell
line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
adaptor sequence: 5' GGAATCGGCACGAG 3' ~3' adaptor
sequence: 5' CTCGACGTTTTTTTTTTT 3'"

```

Query Match	2.7%;	Score 131.8;	DB 45;	Length 864;
Best Local Similarity	52.1%;	Pred. No. 1.1e-26;		
Matches 320;	Conservative 0;	Mismatches 285;	Indels 9;	Gaps 11;

[illegible]

Db	313	TTTGGCTGTACCCCCACTTGGGAAGGCTGGATAATCAGGGCAAGATCAATGCTTGAGC	372
QY	4107	CCTAAGTGAATAATCCAAAAGAGTGGCTGCAGTGGACTTCCAGAAGACAATGAAAGTC	4166
Db	373	GCTCAGAGCAACAGTGCACAAAGAAATGGCTGCAGGTGACCTGNCACATCAGAGGCCAAGTG	432
QY	4167	ACAGAGTAACACTCTCAGGAGTAATAATCTCTGCTTACCAGCATGTATGTGAAGAGTTC	4226
Db	433	ACAGGAATCATCACCCAGGNGCCCGTAGCTTTTGCCACATCAGTATGTGGCGTCTTAC	492
QY	4227	CTCATCTCCAGCAGTCAAGATGGCCATCAGTGCACCTCTTTTTCAGAAATGGCAAGTA	4286
Db	493	AAGTAGCCACAGTAGATGATGTTGTGCAGTGGACTGTATATGANGAGCAAGGAAGCAGC	552
QY	4287	AAGTTTTTCAGGAAATCAAGACTCCTTACACCTGTGGTGAACCTCTCTAGAACCCACCG	4346
Db	553	AAGTCTTCCAGGGGCACTTGGGGAACACACTCCCCACAGAAGACATCTTCCAGAAACCT	612
QY	4347	TTACTGACTGCTAACCCTTGAATTCACCCCCAGAGTTGGGTGCACCGAATGGCCCTGAGG	4406
Db	613	TTCAATGGCTGCTACGTGCGTGTCTTTCAGTGTCTCTGGCATTAACCGCATCACCCCTGGCG	672
QY	4407	ATGAGGTTCTGGG	4420
Db	673	CTGAGCTGTGGG	686

RESULT 12	LOCUS	DEFINITION
A1173222/c	657 bp	EST
A1173222	657 bp	EST
u63a12.x1	Sugano mouse liver	musculus CDNA clone
IMAGE:1450558	3'	similar to gb:M1697 COAGULATION FACTOR V
PRECUSOR (HUMAN):		mrna sequence.

ACCESSION	A1173322	
NID	93720802	
VERSION	A1173322.1	GI:3720802
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 657)
Matta, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Thersing, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE	The WashU-HHMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	On Jan 17, 1998 this sequence version replaced gi:1900316.

Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:923874
Seq primer: custom primer used
High quality sequence stop: 371.

```

FEATURES
source
Location/Qualifiers
1..657
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/map="11q13"
/clone="IMAGE:1450558"
/clone_lib="Sugano mouse liver mlia"
/sex="female"
/dev_stage="adult"

```


NID 91806798
VERSION AA208846.1 GI:1806798
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 633)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The Washu-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1393984.

CONTACT: Marra M/Mouse EST Project
Washu-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:416192
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 443.

FEATURES

Source
1. 633
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:676488"
/clone_lib="Soares mouse NML"
/tissue_type="Liver"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dt) primer [5',
TGTACCAATCTGAAGTGGAGCGCGCGGCGAATCTTTTCTTTT 3'];
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."

GC COUNT 188 a 153 c 158 g 134 t

Query Match 2.5%; Score 121.4; DB 30; Length 633;
Best Local Similarity 53.7%; Pred. No. 9.2e-24;
Matches 296; Conservative 0; Mismatches 251; Indels 4; Gaps 2;

QY 3652 CTGTGATCAAGGTGGATCTGTGGACCAATGATTAATTCACGGCATCAAGACCCAGGGTG 3711
|||||
Db 45 CTGTGATCCAGGTGGACATGCAAGGAAGTGTAGTCACCGGGATACAAACCCAAAGGTG 104
QY 3712 CCCGTCAGAGTCTCCAGCCTCTACATCTCTCAGTTATCATCATGTATAGTCTGTATG 3771
|||
Db 105 CTAACACTACTAAAGTCTCTTACACGAGGAGTTCCAGTGCGCTTACAGCTCTGACC 164
QY 3772 GGAAGAAGTGGCAGACTTAATCGAGGAATTCAGTGAACCTTAATGCTCTTGGCA 3831
|||
Db 165 AAACCAACTGGCAGATCTTCAGAGGAGGAGGAGGAGAGCGTGTATTTTACTGTGTA 224
QY 3832 ATGTGATTCATCTGGGATAAACAATATTTTAAACCTTCAATTAATGCTCGATACA 3891
|||
Db 225 ATTCAGATGGCTCTACAATAAAGAAGATGCACTTGACCCACCATTTGTGGCTAGATACA 284
QY 3892 TCCGTTGCACCCCACTCATATAGCATTCGACGACACTCTTGCGATGGAGTGGGCT 3951
|||
Db 285 TTAGGATACACCCCAACAAATCTTATATAGACCCACCCCTTGCGCTGAGGCTGACGGCT 344

QY 3952 GTGATTTAATAGTTGCGAGCATGCCATTGGGAATGGAGAGTAAAGCAATATCAGATGCAC 4011
|||||
Db 345 GTGAGGTGAACGGATGTTCACACACACTGGCCCTGGAAGATGAGCGATTCAGACAGC 404
QY 4012 AGATTACTGCTTCATCTCTACTTTACCAATATGTTGC--CACCTGGTCTCCTTCAAAAG 4068
|||||
Db 405 AAATTACTGCATCTTCATTTTAAAGTCGTGGTGGGAGACTACTGAGAGCCCTCCCTTG 464
QY 4069 CTCGACTTCACCTCCAGAGGAGAGATATGCCCTGAGACCTCAGGTGATATATCAAAAG 4128
|||
Db 465 CCCGCTGAACGCGCCGTGCCGCTG- AACGCTGGCAGCCAGGGCAACACACACAAGC 523
QY 4129 AGTGCTGCAAGTGGACTTCCAGAGACATGAAGTCACAGAGTAATCTACTCAGGGAG 4188
|||||
Db 524 AGTGGTTTACAGTCTGCTCTCAAAATCAAGAGTTACGCGCATCTAAGCAGGGAT 583
QY 4189 TAAATCTCTG 4199
Db 584 GTAAGTCTCTG 594

RESULT 15
AI225600 1006 bp mRNA EST 29-OCT-1998
LOCUS uJ06C05.y1 Sugano mouse liver mlia Mus musculus cDNA clone
DEFINITION IMAGE:1891112 5' similar to gb:M13699 CERULOPLASMIN PRECURSOR
(HUMAN);, mRNA sequence.
ACCESSION AI225600
NID 93808653
VERSION AI225600.1 GI:3808653
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 1006)
Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.

TITLE The Washu-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Jan 17, 1998 this sequence version replaced gi:2044082.

CONTACT: Marra M/Mouse EST Project
Washu-HHMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:975436
Seq primer: custom primer used
High quality sequence stop: 463.

FEATURES

Source
1. 1006
/organism="Mus musculus"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:1891112"
/clone_lib="Sugano mouse liver mlia"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
(CACTGCTG); Site_2: DraIII (CACCATGTC); 1st strand cDNA
was primed with an oligo(dt) primer
[ATGGGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was
ligated to a DraIII adaptor [TGTGGCTTACTGG], digested

and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTCTGCTCTAAAAGCTGCG and 3' end primer CGACCTGCAGCTCGACACA."

BASE COUNT 300 a 198 c 224 g 274 t 10 others
ORIGIN

Query Match 2.5%; Score 120.8; DB 43; Length 1006;
Best Local Similarity 54.9%; Pred. No. 1.7e-23;
Matches 277; Conservative 0; Mismatches 222; Indels 6; Gaps 2;

```
OY 264 TACAAAAGACTCTGTTGTAGATTACGGATCACCTTTTCACATCGCTAAGCAAG 323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 227 TATAAGAGGCCCTTTATTGTAGTACACAGATGGCACCTTTAGTAAGACTATAGACAAA 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    324 CCACCCCTGGATGGGTCTGCTAGCTCTACACCAAGCTGAGTTTATGATACAGTGTG 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    287 CCAGCCTGGCTAGGGGTTTGTAGCCCTGTCAACAAAGCTGAAGTTGAAGATAAGTTAT 346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 384 ATTACACTTAAGAACATGAGCTTCCATCCTGTGAGTCTTCATGCTGTGTATCCTAC 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 347 GTTCACTTAAGAACCTTGCCCTAGATCTACACTTTTCATGCACATGGGGTAACGTAC 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 444 TGGAAAGCTTGTAGGAGCTGATATGATGATCAGACCAAGTCAAGAGAGAAAGAT 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 407 ACCAAGAGATAGAGGAGAGCCGCTACCCGTGACACACACACTGATTTTCAACGGGCTGAT 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 504 GATAAAGTCTTCCCTGTGGAAGCCATACATATGTCTGCAGGTCCTGAAGAAGATGT 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 467 GACAAAGTGTCTCCCGACACAGATATGTATGTGCTGATGCCC--ATGAGCCAAGT 523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 564 CCAATGGCCTTGACCCACTGTGCTTACCTACTATATCTTCTCATGTGACCTGTA 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 524 CCTGGAGAGGAGACAGACANTTGTGTGACCAAGATTTACCACATCCCATGTGATGCTCA 583
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 624 AAAGACTTGAATTCAGGCTCATTTGAGGCCCTACTAGTATGTAAGAAGGAGTCTGCC 683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 584 NAAGATATTCATCAGACTCATAGGACCTCTAATACTCTGTANANAAGTCTCTATAN 643
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 684 AAGGAAA--AGACACAGACCTTGACACAATTTACTACTTTTGTCTGTATTGATGAA 740
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 644 TAGGAAAAAGAGAAAAATATATGACCAAGAATNTGTAATATGTCTGTGTGATGAA 703
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 741 GGGAAAAGTTGGCACTCAGAAACAA 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
    704 AATCTCAGCTGTATCTTGAAGAA 728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: August 18, 1999, 14:42:40
Job time: 789 sec

